REGULATABLE, CATALYTICALLY ACTIVE NUCLEIC ACIDS

FIELD OF THE INVENTION

[0001] This application is a continuation-in-part of United States Serial No. 60/212,097, filed June 15, 2000.

[0002] The present invention relates generally to the field of catalytic nucleic acids and in particular to regulatable, catalytically active nucleic acids that modulate their kinetic parameters in response to the presence of an effector.

BACKGROUND OF THE INVENTION

[0003] Ribozymes are oligonucleotides of RNA that can act like enzymes and are sometimes called RNA enzymes. Generally, they have the ability to behave like an endoribonuclease, catalyzing the cleavage of RNA molecules. The location of the cleavage site is highly sequence specific, approaching the sequence specificity of DNA restriction endonucleases. By varying conditions, ribozymes can also act as polymerases or dephosphorylases.

[0004] Ribozymes were first described in connection with Tetrahymena thermophilia. The Tetrahymena rRNA was shown to contain an intervening sequence (IVS) capable of excising itself out of a large ribosomal RNA precursor. The IVS is a catalytic RNA molecule that mediates self-splicing out of a precursor, whereupon it converts itself into a circular form. The Tetrahymena IVS is more commonly known now as the Group I Intron.

[0005] Regulatable ribozymes have been described, wherein the activity of the ribozyme is regulated by a ligand-binding moiety. Upon binding the ligand, the ribozyme activity on a target RNA is changed. Regulatable ribozymes have only been described for small molecule ligands such as organic or inorganic molecules. Regulatable ribozymes that are controlled by proteins, peptides, or other macro-molecules.

SUMMARY OF THE INVENTION

[0006] The present invention includes a regulatable, catalytically active nucleic acids (RCANA), wherein the catalytic activity of the RCANA is regulated by an effector. The RCANA of the present invention are, therefore, regulatable in that their activity is under the control of a second portion of the RCANA. Just as allosteric protein enzymes undergo a change in their kinetic parameters or of their enzymatic activity in response to interactions with an effector, the catalytic abilities of the RCANA may similarly be modulated by the effector(s). Thus, the present invention is directed to RCANA that transduce molecular recognition into catalysis.

[0007] As will become apparent below, RCANA are more robust than allosteric protein enzymes in several ways: (1) they can be selected *in vitro*, which facilitates the engineering of particular constructs; (2) the levels of catalytic modulation are much greater for RCANA than for protein enzymes; and (3) since RCANA are nucleic acids, they can potentially interact with the genetic machinery in ways that protein molecules may not.

[0008] It should be noted that the methods described herein may include any type of nucleic acid. For example, these methods are not limited to RNA-based RCANA, but also encompass DNA RCANA and RNA or DNA RCANA. Furthermore, the methods can be applied to any catalytic activity the ribozymes are capable of carrying out. For example, the methods are not limited to ligases or splicing reactions, but could also encompass other ribozyme classes. The methods are also not limited to protein or peptide ligands, but also include other molecular species, such as ions, small molecules, organic molecules, metabolites, sugars and carbohydrates, lipids and nucleic acids. The methods may also be extended to effectors that are not molecules, such as heat or light or electromagnetic fields.

Furthermore, the methods are not limited to ligand-induced conformational changes, but could also take into account 'chimeric' catalysts in which residues essential for chemical reactivity were provided by both the nucleic acid and the ligand, in concert.

[0009] The effector may be a peptide, a polypeptide, a polypeptide complex, or a modified polypeptide or peptide. The effector may even be, e.g., an enzyme or even light (such as visible light) or even a magnet. The effector may be activated by a second effector that acts on the first effector (also referred to herein as an effector-effector), which may be an inorganic or an organic molecule. The polypeptide, peptide or polypeptide complex can be either endogenous, i.e., derived from the same cell type as the polynucleotide, or exogenous, i.e., derived from a cell type different than the cell from which the polynucleotide is derived.

[0010] The polypeptide or peptide may be phosphorylated or dephosphorylated. Alternatively, the effector may include a pharmaceutical agent. In some embodiments, the nucleic acid catalyzes a reaction that causes the expression of a target gene to be up-regulated. In other embodiments, the nucleic acid catalyzes a reaction that causes the expression of a target gene to be down-regulated. If desired, the nucleic acid may be used to detect at least one exogenous effector from a library of candidate exogenous effector molecules. In some embodiments, the nucleic acid and the effector form a nucleic acid-effector complex.

[0011] In some embodiments, the kinetic parameters of nucleic acid catalysis are altered in the presence of a supermolecular structure, e.g., a viral particle or a cell wall. The nucleic acid may further include a regulatory element that can recognize a target molecule of interest. The nucleic acid may in addition include a transducer element that transmits information from the regulatory element to the catalytically active region of the nucleic acid.

[0012] The invention also includes a biosensor that includes a solid support on which at least one regulatable, catalytically active nucleic acid is disposed. The kinetic parameters of the nucleic acid on a target vary in response to the interaction of an effector molecule with the nucleic acid. The regulatable, catalytically active nucleic acid may be immobilized on the support and the reaction may be machine-readable. The solid support may include, e.g., a multiwell plate, a surface plasmon resonance sensor. Regulatable, catalytically active nucleic

acid may be covalently or non-covalently immobilized on the solid support. In some embodiments, the catalytic reaction produces a detectable signal. The substrate may include at least 10 regulatable, catalytically active nucleic acids, at least 100 regulatable, catalytically active nucleic acids, at least 1000 regulatable, catalytically active nucleic acids, at least 10,000 regulatable, catalytically active nucleic acids or even at least 100,000 regulatable, catalytically active nucleic acids.

[0013] Protein and Peptide RCANA. The present invention includes RCANA with catalytic activity that is regulated by a protein or peptide. One embodiment of the present invention involves the in vitro selection of RCANA that are regulated by proteins. A selection scheme for RCANA dependent on protein cofactors has been developed.

[0014] This invention allows the selection of protein-dependent RCANA, which are reagents that can be useful in a variety of applications. For example, protein-dependent RCANA can be used: (1) in chips for the acquisition of data about whole proteomes, (2) as *in vitro* diagnostic reagents to detect proteins specific to disease states, such as prostate-specific antigen (PSA) or viral proteins, (3) as sentinels for the detection of biological warfare agents, (4) as elements in cell-based assays or animal models for drug development studies or (5) as regulatory elements in gene therapies, as described herein. Initially, many protein targets may prove refractive to selection. However, many derivatives of the base method can be developed, to deal with novel targets or target classes.

[0015] Modification of Catalytic Residues of RCANA. In one embodiment of this invention, the RCANA is generated by the modification of at least one catalytic residue. One of the unique features of the present selection protocol relative to others that have previously been published is that the present invention randomizes not only a domain that is pendant on the catalytic core, but a portion of the catalytic core itself. Thus, the selection for ligand-dependence not only yields species that bind to ancillary regions of the RCANA, but that may help stabilize the catalytic core of the RCANA.

[0016] Also provided by the invention is a method of isolating a regulatable, catalytically active nucleic acid created by randomizing at least one nucleotide in the catalytic domain of a catalytically active nucleic acid to create a nucleic acid pool. The

nucleic acid pool whose nucleic acids interact with the catalytic target of the catalytic domain are removed. The method further may also include the step of adding an effector to the remaining pool of nucleic acids. In some embodiments, the method may also include the step of adding an effector to the remaining nucleic acids, wherein the effector acts on the nucleic acids to alter the catalytic activities of the nucleic acids. The method may include optionally the step of purifying the isolated nucleic acid, and, if desired the step of sequencing the isolated nucleic acid. In various embodiments, the step of removing the nucleic acids is under high stringency conditions, moderate stringency conditions, or low stringency conditions.

[0017] Automated Selection of RCANA. The invention further includes the automation of *in vitro* selection, and a mechanized system that executes both common and modified *in vitro* selection procedures. Automation facilitates the execution of this procedure, accomplishing in hours-to-days what once necessitated weeks-to-months. Additionally, the mechanized system attends to other technical obstacles not addressed in "common" *in vitro* selection procedure (e.g., specialized robotic manipulation to avoid cross-contamination). The automation methods are generalizable to a number of different types of selections, including selections with DNA or modified RNA, selections for ribozymes and selections for phage-displayed or cell-surface displayed proteins.

[0018] Automating selection greatly diminishes human error in the actual pipetting and biological manipulations. While programming the robot is often not a trivial task, and can be time consuming, automated selection is far faster and more efficient than manual selection. Time is used preparing samples and analyzing data, rather than performing the actual selection. Additionally, automated selection may include real-time monitoring methods (e.g., molecular beacons, TaqMan®) into the selection procedure and software that can make intelligent decisions based on real-time monitoring.

[0019] *In vitro* sensing (or detection) applications. The current invention also provides for the use of RCANA for detection of a wide variety molecular species *in vitro*. For example, RCANA may be anchored to a chip, such as wells in a multi-well plate. Mixtures of analytes and fluorescently tagged substrates are added to each well. Where

cognate effectors are present, the RCANA will covalently attach the fluorescent tags to themselves. Where RCANA have not been activated by effectors, the tagged substrates are washed out of the well. After reaction and washing, the presence and amounts of co-immobilized fluorescent tags are indicative of amounts of ligands that were present during the reaction. The reporter may be a fluorescent tag, but it may also be an enzyme, a magnetic particle, or any number of detectible particles. Additionally, the RCANA may be immobilized on beads, but they could also be directly attached to a solid support via covalent bonds.

[0020] One advantage of this embodiment is that covalent immobilization of reporters (as opposed to non covalent immobilization, as in ELISA assays) allows stringent wash steps to be employed. Additionally, ribozyme ligases have the unique property of being able to transduce effectors into templates that may be amplified, affording an additional boost in signal prior to detection.

[0021] Modified nucleotides may be introduced into the RCANA that substantially stabilize them from degradation in environments such as sera or urine. The analytical methods of the present invention do not rely on binding per se, but only on transient interactions. The present invention requires mere recognition rather that actual binding, providing a potential advantage of RCANA over antibodies. That is, even low affinities are sufficient for activation and subsequent detection, especially if individual immobilized RCANA are examined (i.e., by ligand-dependent immobilization of a quantum dot).

[0022] Expression of RCANA in cells. The RCANAs of the present invention may also be expressed inside cells. The RCANAs of the present invention that are expressed inside a cell are not only responsive to a given effector, but are also able to participate in genetic regulation or responsiveness. In particular, self-splicing introns can splice themselves out of genes in response to exogenous or endogenous effector molecules.

[0023] The present invention includes RCANA constructs that may be inserted into a gene of interest, e.g., a gene targeting expression vector. The RCANA sequence provides gene specific recognition as well as modulation of the RCANA's kinetic parameters. The kinetic parameters of the RCANA vary in response to an effector. Specifically, in the case of

RCANA that performs self splicing in the presence of the effector, the RCANA may splices itself out of the gene in response to the effector to regulate expression of the gene.

[0024] In another aspect, the invention includes a method of modulating expression of a nucleic acid by providing a polynucleotide that is regulated by a peptide. The polynucleotide may be a regulatable, catalytically active polynucleotide, in which the peptide interacts with the polynucleotide to affect its catalytic activity. The polynucleotide is contacted with the peptide, thereby modulating expression of a nucleic acid. The polynucleotide may be provided in a cell, and the cell may be, e.g., provided in vitro or in vivo and may be a prokaryotic cell or a even a eukaryotic cell.

[0025] The present invention also includes an RCANA construct with a regulatable oligonucleotide sequence having a regulatory domain, such that the kinetic parameters of the RCANA on a target gene vary in response to the interaction of an effector with the regulatory domain.

[0026] *In vivo* sensing (or detection) applications. It is possible to activate or repress a reporter gene (e.g., luciferase) containing an engineered intron in response to an endogenous activator. In this way, luciferase-engineered intron constructs may be used to monitor intracellular levels of proteins or small molecules such as cyclic AMP. This method may be used for *in vivo* measurements in both cellular systems, such as cell culture, and in whole organisms, such as animal models. Such applications may be used for high-throughput screening. If a particular intracellular signal (e.g., the production of a tumor repressor) was desired, compound libraries for pharmacophores that induce the signal (the tumor repressor) are screened for activation of the reporter gene. Thus, the information desired is changed or morphed into the detection of glowing cells.

[0027] Gene therapy applications. Similarly, a gene can be activated or repressed in response to an exogenously introduced effector (drug) for gene therapy. The RCANA may be used for gene expression up regulation (increasing production of the gene product) or down regulation (decreasing the production of the gene product). The construct of one embodiment of the present invention provides a DNA oligonucleotide coding for a catalytic domain and effector binding domain. The advantages of the nucleic acid-based technology of

the present invention include, e.g., the ability to continually modulate gene expression with a high degree of sensitivity without additional gene therapy interventions.

[0028] In another aspect, the invention includes a method of modulating expression of a nucleic acid in a cell by providing a polynucleotide that is regulated by an effector, e.g., a peptide. The polynucleotide may be a regulatable, catalytically active polynucleotide, in which the peptide interacts with the polynucleotide to affect its catalytic activity. The polynucleotide is contacted with the peptide, thereby modulating expression of a nucleic acid. The polynucleotide may be provided in a cell, and the cell may be, e.g., provided in vitro or in vivo and may be a prokaryotic cell or a even a eukaryotic cell.

BRIEF DESCRIPTION OF THE DRAWINGS

[0029] For a more complete understanding of the features and advantages of the present invention, reference is now made to the detailed description of the invention along with the accompanying figures in which corresponding numerals in different figures refer to corresponding parts and in which:

Figure 1 is a depiction of the secondary structure of the Group 1 theophylline-dependent (td) intron of bacteriophage T4 (wild type);

Figure 2a is a photograph of a gel showing activation of the GpITh1P6.131 aptamer construct, together with a graphical representation of the gel, of one embodiment of the present invention;

Figure 2b is a photograph of a gel showing activation of GpITh2P6.133 aptamer construct, together with a graphical representation of the gel of one embodiment of the present invention.

Figure 3 is a schematic depiction of an *in vivo* assay system for group I introns of one embodiment of the present invention.

Figure 4a depicts a portion of the P6 region of the Group I ribozyme joined to the anti-theophylline aptamer by a short randomized region to generate a pool of aptazymes of the present invention.

Figure 4b is a schematic depiction of a selection protocol for the Group I P6 Aptazyme Pool of Figure 4a.

Figure 5 is a diagram of one embodiment of the present invention depicting exogenous or endogenous activation of Group I intron splicing;

Figure 6 is a diagram of another embodiment of the present invention depicting a strategy for screening libraries of exogenous activators;

Figure 7 is a diagram of an alternative embodiment of the present invention for screening libraries of exogenous activators;

Figure 8 is a diagram of yet another alternative embodiment of the present invention for screening libraries of exogenous activators;

Figure 9 is a diagram of an embodiment of the present invention for screening for endogenous activators;

Figure 10 is a diagram of an alternative to the embodiment of Figure 9 of the present invention to screen for endogenous activators;

Figure 11 is a diagram of another embodiment of the present invention to screen for compounds that perturb cellular metabolism;

Figure 12 is a diagram of a further embodiment of the present invention that provides a non-invasive readout of metabolic states;

Figure 13 is a diagram of yet a further embodiment of the present invention wherein endogenous suppressors provide a non-invasive readout of multiple metabolic states;

Figure 14 is a schematic depiction of an example of a work surface for automatic selection procedures of one embodiment of the invention;

Figure 15a is an illustration of the LI ligase aptazyme construct of one embodiment of the present invention;

Figure 15b is an illustration of a modified LI ligase aptazyme construct of Figure 15a of the present invention;

Figure 15c is a schematic diagram of a selection protocol of one embodiment of the present invention;

Figure 16 is a schematic diagram of a method to anchor an aptazyme construct of the present invention to a solid support in one embodiment of the present invention;

Figures 17 (a-d) show the L1 ligase was the starting point for pool design;

Figure 18 (a-d) shows the progression of the L1-N50 selections;

Figure 19 (a & b) shows protein-dependent regulatable, catalytically active nucleic acid sequences and structures;

Figure 20 demonstrates the ribozyme activity with inactivated protein samples;

Figure 21 demonstrates an aptamer competition assays;

Figure 22 shows the binding and ligation activity as a function of protein concentration;

Figure 23 is a flow chart of a method for negative and positive selection of RCANA;

Figure 24 shows the progress of the L1-N50 Rev selection;

Figure 25 (a & b) shows the theophylline-dependent *td* group I intron constructs of the present invention;

Figure 26 shows the design of an FMN-dependent *td* nucleic acid intron splicing construct;

Figures 27 (a-c) show the relative growth curves of theophylline-dependent *in vivo* growth;

Figure 28 shows 3-Methylxanthine dependent *in vivo* growth;

Figure 29 (a & b) shows a schematic of ribozyme ligase array;

Figure 30 shows the results of a regulatable, catalytically active ligase array;

Figure 31 shows the titrations of individual allosteric ribozyme ligases.

DETAILED DESCRIPTION OF THE INVENTION

[0030] While the making and using of various embodiments of the present invention are discussed in detail below, the present invention provides many applicable inventive concepts that may be embodied in a wide variety of specific contexts. The specific embodiments discussed herein are merely illustrative of specific ways to make and use the invention and do not delimit the scope of the invention.

[0031] The present invention includes compositions or matter, methods and automation that permit the creation, isolation, identification, characterization and optimization of regulatable catalytically active nucleic acids. Furthermore, it includes methods to use RCANA for *in vitro* sensing (or detection), *in vivo* sensing (or detection), and gene therapy. Regulatable, catalytically active nucleic acids selected by the method of the present invention also have advantages over other biopolymers that might be used for sensing or gene regulation. Regulatable, catalytically active nucleic acids are more robust than allosteric protein enzymes in several ways: (1) they can be selected *in vitro* (facilitating the engineering of particular constructs); (2) the levels of catalytic modulation are much greater than those typically observed with protein enzymes; and (3) since regulatable, catalytically

active nucleic acids are nucleic acids, they can potentially interact with the genetic machinery in ways that protein molecules may not.

[0032] The method is not limited to RNA pools, but may also encompass DNA pools or modified RNA pools. Modified nucleotides may be introduced into the regulatable, catalytically active nucleic acids that substantially stabilize them from degradation in environments such as sera or urine. The method is not limited to ligases, but could also encompass other ribozyme classes. The method is not limited to protein-induced conformational changes, but could also take into account 'chimeric' catalysts in which residues essential for chemical reactivity were provided by both the nucleic acid and the protein in concert. Initially, many protein targets may prove refractive to selection. Many derivatives of the base method can be developed, however, to deal with novel targets or target classes.

A. Protein Dependent RCANA

[0033] Effector-dependent ribozymes have been shown to be responsive to small organic compounds, such as ATP and theophylline. The present inventors recognized the need for effector-dependent ribozymes, or as used herein, "regulatable, catalytically active nucleic acids" that are responsive to larger molecules, such as, e.g., peptides or proteins. The peptides, proteins or other large molecules may be provided from endogenous sources (e.g., expressed within a cell or cell extract), or exogenous sources (added or expressed in a cell or cell extract).

[0034] In order to understand the present invention, it is important to understand that previous attempts to make catalytically active nucleic acids that interact and respond to a large effector, by the inventors and others, have failed. Initially, attempts were made to generate protein-dependent ribozymes by the addition of aptamers (known binding sequences) to ribozymes (catalytically active domains). This design and method was unsuccessful in providing regulatable nucleic acids. Next, attempts were made to generate protein-dependent ribozymes by adding random sequence regions between an aptamer (binding) and a ribozyme (catalytic) and selecting for effector-dependence. These attempts

were also unsuccessful. Next, the inventors attempted to generate protein-dependent ribozymes by adding a large random sequence region to the catalytic cores of ribozymes and selecting for effector-dependence. These attempts were also unsuccessful. In other words, all previously detailed methods for the generation of ribozymes that were dependent on small organic compounds were unsuccessful for generating ribozymes that were dependent on proteins.

[0035] To date, the present inventors have selected a number of protein- and peptide-dependent ribozyme ligases. One example is the isolation of a protein-dependent, regulatable, catalytically active nucleic acid with an activity that was increased in a standard assay by 75,000-fold in the presence of its cognate protein effector, tyrosyl tRNA synthetase from *Neurospora* mitochondria (Cyt18). The Cyt18-dependent ribozyme was not activated by non-cognate proteins, including other tRNA synthetases.

[0036] A protein-dependent, regulatable, catalytically active nucleic acid was also created and selected with an activity that was increased by 3,500-fold in the presence of its cognate protein effector, hen egg white lysozyme. The lysozyme-dependent ribozyme was not activated by most non-cognate proteins, including T4 lysozyme, but was activated by a very closely related protein, turkey egg white lysozyme. Moreover, the protein-dependent ribozyme was inhibited by a RNA binding species that specifically bound to lysozyme. In other words, the activation of these protein-dependent ribozymes was highly specific.

[0037] A peptide-dependent, regulatable, catalytically active nucleic acids was also created and isolated with activity was increased by 18,000-fold in the presence of its cognate peptide effector, the arginine-rich motif (ARM) from the HIV-1 Rev protein. The Rev-dependent nucleic acid was not activated by other ARMs from other viral proteins, such as HTLV-I Rex. Using the present invention, regulatable, catalytically active nucleic acids may be developed that are regulated by any of a vast number of proteins.

[0038] As will be clear from the continued description, protein dependent RCANAs are useful in a variety of applications. For example, protein-regulated catalytically active nucleic acids can be used (1) for the acquisition of data about whole proteomes, (2) as *in vitro* diagnostic reagents to detect proteins specific to disease states, such as prostate-specific

antigen or viral proteins, (3) as sentinels for the detection of biological warfare agents, or (4) as regulatory elements in gene therapies.

B. Modification of Residues in Catalytic Domain

[0039] In one embodiment, the present invention randomizes a portion of the catalytic core itself, not necessarily a domain that is pendant on the catalytic core. One example for selection using the present invention was using the L1 ligase. The catalytic core of the L1 ligase has been mapped by deletion analysis and by partial randomization and re-selection. Figure 15a depicts the LI ligase that was the starting point for pool design. Stems A, B, and C are indicated. The shaded region contains the catalytic core and ligation junction. Primer binding sites are shown in lower case, an oligonucleotide effector required for activity is shown in italics, and the ligation substrate is bolded. The 'tag' on the ligation substrate can be varied, but was biotin in the exemplary selection described herein. The LI pool contains 50 random sequence positions and overlaps with a portion of the ribozyme core. In Figure l5b, Stem B was reduced in size and terminated with a stable GNRA tetraloop, but stem A was unchanged.

[0040] A pool was synthesized in which the random sequence region spanned the catalytic core. Protein-dependent ribozymes were selected from this random sequence pool by selecting for the ability to ligate an oligonucleotide tag in the presence of a protein effector followed by capturing the oligonucleotide tag on an affinity matrix, followed by amplification *in vitro* or *in vivo*. Because the catalytic core has been randomized, the selection for protein-dependence not only yields species that may bind to ancillary regions of the ribozyme, but species in which the protein effector actually helps to organize the catalytic core of the ribozyme.

[0041] Selection for protein-dependence from a pool in which at least a portion of the catalytic core of the ribozyme is randomized differs from selection for protein-dependence from a pool in which the catalytic core is not randomized. For example, the catalytic core of the protein-dependent ribozymes that was selected differed substantially from the catalytic

core of the original ribozyme and the catalytic core of other, non-protein-dependent ribozymes selected based on the original ribozyme.

[0042] Figure 15a depicts the LI ligase that was the starting point for pool design in the Cytl8 RCANA selection, as an example of a protein-activated regulated, catalytically active nucleic acid. Stems A, B, and C are indicated. The shaded region contains the catalytic core and ligation junction. Primer binding sites are shown in lower case, an oligonucleotide effector required for activity is shown in italics, and the ligation substrate is bolded. The 'tag' on the ligation substrate can be varied, but was biotin in the exemplary selection described herein. The LI pool contains 50 random sequence positions and overlaps with a portion of the ribozyme core. In Figure l5b, Stem B was reduced in size and terminated with a stable GNRA tetraloop, but stem A was unchanged.

[0043] Because one or more residues in the catalytic core have been randomized, the effectors may add essential catalytic residues for a given reaction. That is, both the effector molecule and the regulatable, catalytically active nucleic acids contribute a portion of the active site of the ribozyme. For example, using the method of the present invention a ribozyme and an effector molecule that would only carry-out poorly an enzymatic function independently may perform that enzymatic function upon interaction with one another. As such, a regulatable, catalytically active nucleic acid that contributes a guanosine and an adenosine and a protein effector that contributes a histidine together form a complex that has greater activity than either of the individual compounds. Using the methods disclosed herein it is possible to identify a chimeric effector:ribozyme (e.g., a protein:RNA complex) active site that would lead to catalysis. The invention describes ribozymes that have a detectable, basal chemical reactivity, and that the presence of the effector modulates this basal chemical reactivity. It is for this reason that the present invention differs significantly from other inventions which have claimed protein:RNA complexes in which no basal catalytic activity exists in the ribozyme or protein alone.

C. Selection of RCANA

[0044] Figure 15c schematically shows the following selection scheme: the RNA pool was incubated with a biotinylated tag and reactive variants were removed from the population. The remaining species were again incubated with a biotinylated tag in the presence of the target (for example the protein Cytl8). Reactive variants were removed from the population and preferentially amplified by reverse transcription, PCR, and *in vitro* transcription.

[0045] Ligand-dependent, regulatable, catalytically active nucleic acids selected by this method differ from functional nucleic acids selected from random sequence pools. Selection for ligand-dependence requires a selection for catalytic activity as opposed to a selection for binding. Therefore, protein-dependent, regulatable, catalytically active nucleic acids are not aptamers. The composition of matter of a selected protein-dependent ribozyme will be different than the composition of matter of a selected aptamer. For example, the sequence of the lysozyme-dependent ribozyme is different from the sequence of anti-lysozyme aptamers. An important feature of the present invention is that the regulatable, catalytically active nucleic acids disclosed herein only required recognition rather than selected or enhanced binding ability. For example, the affinity of lysozyme for the naïve, unselected RNA pool is identical to the affinity of lysozyme for the selected, regulatable, catalytically active nucleic acid. The only difference is that the way in which lysozyme is recognized by the regulatable catalytically active nucleic acids leads to activation, while for the pool as a whole non-specific binding does not lead to activation. In other words, binding is a concomitant but secondary function of selection for regulatability; that is, the regulatable ribozymes disclosed herein may bind the effector or target very poorly, but upon interaction the activity of the ribozyme may nonetheless be modulated.

D. Automated Selection of RCANA

[0046] Robotic workstations have become essential to high-throughput manipulations of biomolecules, such as in high-throughput screening for drugs with a particular mechanism of action. The invention also includes the automation of *in vitro* selection procedures, and a mechanized system that executes both common and modified *in vitro* selection procedures. Automation facilitates the execution of this procedure, accomplishing in hours to days what

once necessitated weeks to months. In particular, the present inventors have adapted a robotic workstation to the selection of aptamers and ribozymes. However, the automation methods are generalizable to a number of different types of selections, including selections with DNA or modified RNA, selections for ribozymes, and selections for phage-displayed or cell-surface proteins.

[0047] In short, *in vitro* selection involves several components: generation of a random sequence pool, sieving the random sequence pool for nucleic acid species that bind a given target or catalyze a given reaction, amplification of the sieved species by a combination of reverse transcription, the polymerase chain reaction, and *in vitro* transcription. Beyond the generation of the random sequence pool, each of these steps can potentially be carried out by a robotic workstation. The pool can be pipetted together with a target molecule. If the target is immobilized on a magnetic bead, then the nucleic acid:target complex can be sieved from solution using an integrated magnetic bead collector. Finally, selected nucleic acid species can be eluted from the complex and amplified via a series of enzymatic steps that include the polymerase chain reaction carried out via an integrated thermal cycler.

[0048] There are many potential ways in which binding species can be sieved from a random sequence population. However, not all of these methods are amenable to automated selection. For example, to select aptamers, others have suggested that targets can be immobilized onto microtitre plates and binding species can be sieved by panning. The present inventors have had little success with this method, likely because panning is a relatively inefficient, low stringency method for sieving. Instead, the present inventors have discovered that when targets are immobilized on beads and mixed with a random sequence pool, binding species can be efficiently sieved from non-binding species by filtration of the beads. Beads can be readily manipulated by pipetting, allowing for the facile recovery and elution of the binding species, which are then amplified and carried into subsequent rounds of selection. This method differs from the magnetic bead capture method, and can be carried out with much higher stringency. This method is novel, and has not previously been used for *in vitro* selection experiments.

[0049] Figure 14 depicts schematically an exemplary work surface for yet another embodiment of the present invention: automated selection. See, J.C. Cox, et al., Automated RNA Selection Biotechnol. Prog., 14, 845 850, 1998.

[0050] Base protocol. Automated selection involves several, sequential automated steps. Several modules are placed on the robotic work surface, including a magnetic bead separator, and enzyme cooler, and a thermal cycler. After manually preparing reagents and preloading those reagents (including random pool RNA, buffers, enzymes, streptavidin magnetic beads, and biotinylated target) and tips onto the robot, a program is run. The selection process, automated by the robot, goes as follows: RNA pool is incubated in the presence, of biotinylated target conjugated to streptavidin magnetic beads. After incubation, the magnets on the magnetic bead separator are raised, and the beads (now bound by pool RNA - the selected nucleic acids) are pulled out of solution. Thus, the beads can be washed, leaving only RNA bound to targets attached to beads. These RNA molecules are reverse transcribed, reamplified via PCR, and the PCR DNA is *in vitro* transcribed into RNA to be used in iterative rounds of selection.

[0051] The Bioworks method for *in vitro* selection. This scripted programming method contains all movements necessary in order to facilitate automated selection. This includes all physical movements to be coordinated, and also communication statements. For instance, five rounds of automated selection against a single target requires over 5,000 separate movements in x, y, z, t coordinate space. Additionally, the method also holds all relevant measurements, offsets, and integrated equipment data necessary to prevent physical collisions and permit concerted communication between devices.

[0052] "Beads on filter" selections. While the vast majority of manual selections have been performed on nitrocellulose-based filters, a small few have also been performed on solid surfaces, such as beads. A novel selection scheme was developed whereby selection is performed on magnetic beads that are placed on nitrocellulose filters, and washed as the bead is the selection target itself. This method allows for much greater specificity of selection, thereby promoting 'winning' molecules to amplify in greater number, and thus reduce the overall amount of rounds necessary to complete the selection procedure. Manual selection

does not involve a combination of surfaces to enhance selection. An alternative method is to take the magnetic beads, or nucleic acids attached to beads using methods other than beads, and running buffer over the beads and through a filter. It has been found that a complete filter washing step provides improved performance in the selection due to decreased background. One example of the automation of such a methods would be to remove, e.g., nucleic acids attached to the beads by placing the beads in a 96-well plate with a filtered bottom, the beads washed with buffer followed by subsequent elution of the target nucleic acids.

[0053] Cross-contamination avoidance. The introduction of contaminating species of nucleic acid strands in a manual selection may be commonplace. This is especially true if selection is done against multiple targets in parallel, and also when a researcher reuses the same pool for different selection tools. Contaminating species have been shown in the past to interfere with a manual selection such that it could not be completed. Automated *in vitro* selection takes steps to minimizing and/or eliminate the possibility of cross-contamination between pools and targets. Movement of the mechanical pod along the work surface is unidirectional when carrying potentially contaminating material. This movement away from 'clean' things and only towards items that have already been exposed to replicons greatly diminishes the possibility of cross-contaminating reactions. The only circumstance in which the pod reverses its direction is to acquire a new, clean pipette tip. Additionally, the reagent trays were sealed with aluminum foil for a physical barrier between the environment and unexposed reagents. See Figure 14, a layout of the robotic work surface that reduces cross-contamination.

[0054] Using this method the present inventors have successfully selected aptamers against a number of protein targets, including Cyt18, lysozyme, the signaling kinase MEK1, Rho from a thermophilic bacteria, and the Herpes virus US11 protein. The robot can perform 6 rounds of selection / day versus individual protein targets, and selections are typically completed within 12 - 18 rounds. In each instance, selected populations showed a substantially greater affinity for their cognate proteins than naïve populations. In addition, when selected populations were sequenced one or more sequence families typically

predominated. Sequence families are a hallmark of a successful selection, and indicate that the robotic method faithfully recapitulates manual selection methods.

[0055] The use of beads for target immobilization allows automated selection to be generalized to virtually any target class. For example, small organic molecules could be directly conjugated to beads. Similarly, antibodies could be conjugated to beads and in turn could be used to capture macromolecular structures, such as viruses or cells.

[0056] In another embodiment, the robotic workstation can be used for the selection of nucleic acid catalysts. For example, a DNA library was incubated that contained an iodine leaving group at its 5' end with a DNA oligonucleotide substrate containing a 3' phosphorothioate nucleophile and a 5' biotin. The biotin can be captured on beads bearing streptavidin, and the beads can in turn be captured either by magnetic separation or by filtration. Any molecules in the DNA pool that ligate themselves to the biotinylated substrate are co-immobilized with that substrate. Immobilized species can be directly amplified following transfer to the integrated thermal cycler. The inclusion of a biotin on one of the primers used for amplification allows single-stranded DNA to be prepared by denaturation of the non-biotinylated strand in base, followed by neutralization of the solution. While this method has proved successful for the selection of deoxyribozyme ligases, variations could also have been attempted. For example, the biotinylated DNA oligonucleotide substrate could have been pre-immobilized on beads, and the DNA pool incubated with the beads. In this instance, any molecules in the DNA pool that ligate themselves to the substrate will also be directly captured on the beads.

[0057] The use of beads for catalyst immobilization immediately suggests other selection protocols. For example, nucleic acid cleavases could be selected by first immobilizing a pool on the beads, then selecting for those species that cleave themselves away from the beads. Similarly, nucleic acid Diels-Alder synthetases may be selected by first immobilizing a diene on the beads, creating a nucleic acid pool that terminates in a dienophile, and selecting for those species that most efficiently conjugate the diene and dienophile.

[0058] This method can be applied to the selection of RCANAs. The ability to use a robotic workstation to select for ligases demonstrates that it is possible to select for regulatable ribozymes. For example, the selection protocols described in this invention can be altered so that ligases that immobilized themselves in the absence of a protein effector are removed from the random sequence population, while ligases that subsequently immobilized themselves once a protein effector were added are transferred to the integrated thermal cycler, amplified, and used for additional rounds of selection. This automated selection methods for regulatable ribozymes can readily be extended to other classes or catalysts than ligases, such as cleavases or Diels Alder synthetases by those skilled in the art.

[0059] Automating selection greatly diminishes human error in the actual pipetting and biological manipulations. While programming the robot is often not a trivial task, and can be time-consuming, automated selection is far faster and more efficient than manual selection. The scientist's time is thus put to better use preparing samples and analyzing data, rather than performing the actual selection. Additionally, automated selection may include real-time monitoring methods (e.g., molecular beacons, TaqMan) and software that can make intelligent decisions based on real-time monitoring.

E. Chip-based RCANA for *in vitro* detection applications

[0060] Regulatable catalytically active nucleic acids are especially useful for biosensor applications. For example, different protein-regulated catalytically active nucleic acids may be anchored to a surface, such as wells in a multi-well plate. Mixtures of analytes and fluorescently tagged substrates are added to each well. Where cognate effectors are present, the protein-regulated catalytically active nucleic acids will covalently attach the fluorescent tags to themselves. Where protein-regulated catalytically active nucleic acids have not been activated by effectors, the tagged substrates will be washed out of the well. After reaction and washing, the presence and amounts of co-immobilized fluorescent tags are indicative of amounts of ligands that were present during the reaction.

[0061] In one embodiment of the invention, the reporter may be a fluorescent tag, but it may also be an enzyme, a magnetic particle, or any number of detectable particles.

Additionally, the protein-regulated catalytically active nucleic acids may be attached to beads or non-covalently linked to a surface rather than covalently linked to a surface.

[0062] One advantage of this method is that covalent immobilization of reporters (as opposed to non-covalent immobilization, as in ELISA assays) allows stringent wash steps to be employed. Additionally, ribozyme ligases have the unique property of being able to transduce effectors into nucleic acid templates that can be amplified, affording an additional boost in signal prior to detection.

[0063] Another advantage is that the analytical methods of the present invention do not rely on binding per se, but only on transient interactions. The present invention requires mere recognition rather than a binding event that must be physically isolated, providing a potential advantage of protein-regulated catalytically active nucleic acids over antibodies. That is, even low affinities are sufficient for activation and subsequent detection, especially if individual, immobilized protein-regulated catalytically active nucleic acids are examined (i.e., by ligand-dependent immobilization of a quantum dot).

[0064] Figure 16 schematically depicts one way to anchor aptazymes to a chip for a particular embodiment of the present invention. In this schematic, different ribozyme ligases (indicated by different colored allosteric sites) are shown immobilized on beads in wells, and mixtures of analytes (differentiated by shape and color) and fluorescently tagged substrates have been added to each well. In the middle panel of this figure, where cognate effectors are present (same color analyte and allosteric site), the aptazymes will covalently attach the fluorescent tags to themselves. Where RCANA have not been activated by effectors, the tagged substrates are washed out of the well. In the last panel of Figure 16, after reaction and washing, the presence and amounts of co-immobilized fluorescent tags are indicative of amounts of ligands that were present during the reaction.

[0065] In the embodiment of Figure 16, the reporter may be a fluorescent tag, but it may also be an enzyme, a magnetic particle, or any number of detectible particles. Additionally, the RCANA could be immobilized on beads, but they could also be directly attached to a solid support via covalent bonds.

[0066] One advantage of this embodiment is that covalent immobilization of reporters allows stringent wash steps to be employed. This can be distinguished from to non covalent immobilization assays such as ELISA assays where stringent washing may destroy the signal. An additional advantage is that ribozyme ligases have the unique property of being able to transduce effectors into templates that can be amplified, affording an additional boost the in signal prior to detection.

[0067] Additionally, the method of the present invention contemplates that the RCANA construct may be amplified by polymerase chain reaction. Finally, the method contemplates that the RCANA oligonucleotide sequence of the construct may include RNA nucleotides, so that the method further includes reverse transcription of the RNA using reverse transcriptase to produce a DNA complementary to the RNA template.

[0068] Modified nucleotides may be introduced into the RCANA that substantially stabilize them from degradation in environments such as sera or urine. The analytical methods of the present invention do not rely on binding per se, but only on transient interactions. The present invention requires mere recognition rather that actual binding, thus providing a potential advantage of RCANA over antibodies. That is, even low affinities are sufficient for activation and subsequent detection, especially if individual immobilized RCANA are examined (i.e., by ligand-dependent immobilization of a quantum dot).

F. *In vitro* engineering and selection of RCANAs for *in vivo* applications

[0069] The above discussion has disclosed methods for the *in vitro* creation of RCANAs, and has disclosed some of their *in vitro* applications. In the following section we describe the design, engineering, and *in vitro* selection of RCANAs for *in vivo* applications.

[0070] This invention utilizes ribozymes that can alter the level of mRNAs in a cellular system. In one embodiment, the ribozyme can be a self splicing intron, such as the group I intron. This ribozyme can be inserted into a gene. If the ribozyme is active, it will catalyze the a self-splicing reaction that removes itself from the gene, allowing accurate expression of the gene. In another embodiment, the ribozyme may be one that acts in trans to cleave a mRNA. Again, changing the activity of the ribozyme will lead to a change in the

level of the mRNA in the cell, thereby altering the level of the protein coded by that gene. Those skilled in the art will recognize that other ribozyme activities may be used. For the purpose of illustration, the invention is now described in detail with the use of the self splicing intron.

[0071] The intron is first modified to function as an RCANA. Briefly, the methods described above can be used generate RCANA introns. A pool of potential RCANA introns is created by randomizing one or more regions of the intron. The randomized region optionally includes one or more residues from the catalytic core. A selection protocol is then developed that allows the active RCANA introns to be partitioned from the inactive ones. For example, the active RCANA introns can be partitioned from the inactive RCANA intron based on the mobility in gel electrophoresis. Other methods will be clear to those skilled in the art. Based on this partitioning method, an iterative procedure of partitioning and subsequent amplification of the RCANA introns is used to select RCANAs that are regulated by an effector. With the exception of the partitioning method, this procedure is essentially identical to the selection described about for RCANA ligases.

[0072] As an alternative to the selection of RCANA introns, it is also possible to engineer RCANA introns. For example, one of the stem-loop structures of the intron can be replaced by an aptamer for the desired effector. Interaction of the effector with this engineered RCANA intron-will result in a modulation of the RCANA intron activity. Because an aptamer is different from an regulatory element (as was detailed above), the present method will, in general, lead to RCANAs that are regulated by the effector. However, as will be shown in an example below, an important aspect of the current invention is that this level of regulation can be adequate for *in vivo* applications.

G. *In vivo* selection and optimization of RCANAs.

[0073] Here we disclose methods to generate RCANAs by using *in vivo* selection. Figure 4b shows a selection protocol for the Group I P6 RCANA Pool of Fig. 4a. Positive and negative selections are made *in vitro* to select Group I RCANA that are dependent on activator. The selections are described above in Example 2 for a specific embodiment of the

present invention – a theophylline dependent RCANA. *In vivo* screens and selections are used to select Group I RCANA that exhibit strong theophylline-dependence. The selected RCANA are mixed at various ratios with mutant Group I ribozymes that splice at a low but continuous level to determine the level at which RCANA can be selected against background. Because activation domains are often in the form of a stem-loop, the mutations can be concentrated in a single stem loop structure of the RCANA intron. In an alternate embodiment, the mutations can include catalytic residues. In yet another embodiment, the mutations are randomly dispersed in the intron. Finally, the best theophylline-dependent Group I aptazymes that have been derived by any of the methods described herein may undergo further selection by partially randomizing their sequences and selecting for improved *in vivo* performance.

[0074] Strategies similar to those depicted in Figs 4a and 4b may be used to develop RCANA on any desired effector. Positive and negative *in vitro* selection such as depicted in Figure 4b are described above in Example 2 for a specific embodiment of the present invention.

[0075] From 10^6 to 10^{10} variants can be efficiently transformed as described herein, sufficient to encompass most variants in the populations discussed so far. This efficiency of transformation, however, is likely to be insufficient to encompass a significant fraction of a completely random pool. Nonetheless, sequences have been selected from completely random expressed pools that can protect bacteria from bacteriophage infection.

[0076] The above procedure described how to select *in vivo* RCANAs. A similar procedure can be used to optimize engineered RCANAs. Residues in the RCANA that might include the ligand binding region, structural stem-loops, or even catalytic residues can be mutated. The selection procedure described above is then used to select for optimized RCANAs.

[0077] Finally, since the rules that govern Group I intron splicing in different gene contexts are well known to those skilled in the art, the skilled artisan can remove RCANA introns from one context and modularly insert them into other genes. Should the efficiency

or effector-dependence of intron splicing be compromised in the new gene, the intron may be reaccommodated to its new genetic environment by a selectable marker to the interrupted gene of interest and selecting for an effector-dependent phenotype.

[0078] To the extent that Group I aptazymes are self-sufficient, they should also function in eukaryotic cells, including human cells. Selecting for effector-dependence may also be performed in eukaryotic cells. Selection in eukaryotic systems may be performed, e.g., by fusing the gene of interest to a reporter gene such as GFP or luciferase. Variants of the RCANA that promote effector-dependent protein production may then be selected using a FACS. A pool of 10^6 to 10^8 variants may be screened by this procedure, a range comparable to the bacterial system previously described.

H. *In vivo* detection applications

[0079] Using the present invention, it is possible to activate or repress a reporter gene (e.g., luciferase or GFP) containing an engineered intron in response to an endogenous protein activator, or a post-translationally modified form of an endogenous protein activator (e.g., protein kinases such as ERK 1 and phosphorylated ERK 1). It is also possible to activate or repress a reporter gene (e.g., luciferase or GFP) containing an engineered intron in response to small molecule effectors (e.g., cyclic AMP, glucose, bioactive peptides, bioactive nucleic acids, or low molecular weight drugs such as antibiotics, antineoplastics or the like.). Thus, reporter gene-engineered intron constructs may be used to monitor intracellular levels of proteins, post-translationally modified forms of proteins or small molecules such as cyclic AMP and the like. Such applications may be used for high-throughput cell-based assays and screens for drug leads or for drug optimization and development.

[0080] Bacterial strains such as E. coli, and B. subtilis, or yeast strains such as S. cerevisiae, and S. pombe may be transformed with an expression vector encoding a reporter gene regulated by an intron RCANA, and these engineered microbial cell lines may be used for cell-based assays and tests for drug discovery and development. Similarly, standard mammalian cell lines such as CHO, NIH3T3, 293, and 293T may be transfected with appropriate vectors (e.g., pCDNA, pCMV, or retrovirus), that are engineered to contain

RCANA-regulatable reporter genes, and these re-engineered cell lines may be used subsequently for cell-based assays and tests. In another use of the RCANA reporter gene technology, tumorigenic cell lines such as LNCaP, MCF-7, MDA-MB-435, SK-Mel, DL1, PC3, T47D and the like, may be transfected *in vitro* with appropriate vectors encoding an RCANA-regulatable reporter gene. These re-engineered tumorigenic cell lines may be used in cell-based screens for the discovery and development anti-neoplastic drugs.

[0081] In another *in vivo* application, reporter gene – intron RCANA constructs (e.g., luciferase or GFP) may be used to generate live animal models for use in drug development. In one embodiment the RCANA-intron construct may be used in an engineered tumorigenic cell line to indicate the levels of a target molecule used to generate a tumor xenograft in nude mice. Mice bearing the tumors derived from the engineered cell line may then be used to screen for drugs that alter the level of the target molecule. For example, a transfected MDA-MB-435 line engineered to express a GFP gene under regulatable control by intron response to the protein activator phospho-ERK 1 is used to screen for drugs which both inhibit tumor growth and block formation of phospho-ERK. In another embodiment of the RCANA intron invention, transgenic mouse models may be generated in which tissue or cell type specific expression of the reporter gene is controlled by the effector activated RCANA intron. For example transgenic mice expressing a phospho-VEGF receptor tyrosine kinase (RTK) specific RCANA regulated GFP gene under control of the MMTV (mouse mammary tumor virus) promoter would show expression of GFP in mouse mammary tissue in a phospho-VEGF RTK dependent manner. Furthermore, these mice may be used to screen compounds for anti-VEGF RTK activity.

[0082] Figure 5 is a diagrammatic representation of another embodiment of the present invention. Exogenous or endogenous activation of Group I intron splicing is depicted. A reporter gene such as Luciferase or beta-Gal is fused to the gene of interest which also contains the group I intron (td). Splicing-out of the Group I intron is induced by an effector, shown in the diagram as a protein, in this case Cyt18, by the shaded oval. Activation of the RCANA and auto-excision of the intron results in expression of the reporter gene to detect the desired reaction. The use of a reporter gene of this embodiment may be suitable for use in eukaryotic systems.

[0083] Figure 6 is a diagram of another embodiment of the present invention. Libraries of candidate exogenous activators (E_{1-n}) may be generated from a randomized RCANA pool indicated by the triangle. As in the embodiment of Figure 5, a reporter gene is expressed in cells where the exogenous activator activates the RCANA to release the intron from the gene. As will be known to those of skill in the art any number of current and future libraries may be used with the present invention.

[0084] Figure 7 depicts an alternative embodiment for screening libraries of exogenous activators. In the embodiment of the present invention of Figure 7, Group I introns are induced into trans-splicing. Extracted and amplified introns are used to transform cells.

[0085] Figure 8 shows yet another alternative embodiment for screening libraries of exogenous activators of the present invention. In the embodiment of Fig. 8, the effector (shaded oval), shown in this illustration as protein Cyt18, is mutagenized (triangle) to form an effector library. A second effector (E1-n) interacts with and activates one or more members of the effector library. The effector-effector complex is exposed to the gene containing both the Group I intron and a reporter gene. Cell sorting reveals the cells that express the reporter gene to indicate successful activation of the RCANA by the effector-effector complex.

[0086] Figure 9 is a diagram of an embodiment of the present invention for screening for endogenous activators. In this embodiment, an endogenous effector, in this illustration shown as a protein activator from endogenous or transformed origin (shaded oval), activates self-splicing of the Group I intron. Cell sorting is used to reveal the expression of the reporter gene. To protect against spontaneous auto-excision of the intron, the gene may be transferred into a different background system such as yeast or E. coli, for example.

[0087] Figure 10 depicts an alternative to the embodiment of Figure 9 to screen for endogenous activators of the present invention. In Figure 10, the activator that is being screened for may include, inter alia, a phosphorylated protein, a product of ubiquitination, or a protein-protein complex. For example, a protein activator, shown as the small shaded oval, may phosphorylate an effector such as Cyt18, shown as a large shaded oval with the

phosphorylation indicated by the shaded rectangle. The phosphorylated effector activates intron self-splicing with concomitant expression of the reporter gene, shown here for illustration as Luciferase or beta-Galactosidase.

[0088] Figure 11 shows yet another embodiment of the present invention to monitor compounds that perturb cellular metabolism. In this embodiment, a ribozyme similar to described in Figure 6, and designated in this diagram by a line with a triangle is activated by a protein effector, shown as a shaded oval in Figure 11. The protein effector may be a phosphoprotein, an induced protein, or a protein complex, for example. One or more second effectors, designated as a series of circles, alters the level of or degree of modification of the protein effector. The source of the second effectors may be endogenous or the effectors may be the product of a transformation construct used to transform a cell. Alteration of the level or modification of the protein effector results in an alteration in the expression of the reporter gene (shown as a dark circle with "lightning bolts"). The functioning of the gene of interest may thereby be perturbed, providing information about the function and/or regulation of the gene or gene product. Figure 11 describes a method for taking the products of the screen described in Figures 8 and 10 and using them to monitor cellular or metabolic states.

[0089] Figure 12 shows a further embodiment of the present invention that provides a non-invasive readout of metabolic states. An RCANA construct of the present invention may be introduced to a gene of interest. A protein suppressor from either an endogenous source from the product of cell transformation activates self-splicing of the RCANA, leading to expression of the endogenous gene, shown here again as a dark circle with lightning bolts. Whether or not the gene of interest is expressed upon release of the RCANA intron from the gene provides information about the metabolic state of the gene of interest. The embodiment of the present invention of Figure 12 thus provides a non-invasive means to determine the metabolic state of an organism with regard to a gene of interest.

[0090] Figure 13 depicts a further embodiment of the present invention wherein endogenous suppressors provide a non-invasive readout of multiple metabolic states. Multiple protein activators (endogenous or transformed) are exposed to a pool of Group I introns of the present invention. The pool comprises introns with length polymorphisms that

are depicted in Figure 13 by a discontinuity or break in the line representing the Group I intron (thick line) residing in a gene of interest (thin line). Activation of the RCANA leads to trans-splicing among the various polymorphisms. The products of trans-splicing may be extracted and amplified. Separation of the trans-splicing products by gel electrophoresis provides a read out of the protein function or the metabolic pathway. The readout may even be digitized for analysis.

I. *In vivo* uses of RCANAs for Gene Therapy

[0091] One important feature of using RCANAs and the method of the present invention for gene therapies is that regulated introns may be used to control gene expression, for any of a variety of genes, since the introns may be inserted into and be engineered to accommodate virtually any gene. Moreover, since the RCANAs may be engineered to respond to any of a variety of effectors, the characteristics of the effector (oral availability, synthetic accessibility, pharmacokinetic properties) may be chosen in advance. The drug is chosen prior to engineering the target of the drug. In part because of these extraordinary capabilities, RCANA provide perhaps the only viable route to medically successful and practical gene therapies. Drugs may be given throughout the treatment (or lifetime) of a patient who had undergone a single initial gene therapy. In addition, by making the gene therapy regulatable, the amount of a gene product may be easily increased or decreased in different individuals at different times during the treatment by increasing or decreasing the doses of effectors.

[0092] The present method also includes transforming a cell with the RCANA construct so that the construct is inserted into a gene of interest. An effector is provided to activate the RCANA so that administering to the cell an effective amount of the effector induces the RCANA to splice itself out of the gene to regulate expression of the gene.

[0093] The method of the present invention contemplates that the RCANA construct may be a plasmid. The method then further includes transforming the cell with the plasmid. The method of the present invention also contemplates ligating the RCANA construct into a vector and transforming the cell with the vector.

DEFINITIONS

[0094] As used herein, the term "regulatable, catalytically active nucleic acid" or "RCANA" means a ribozyme or nucleic acid enzyme that is regulated by an effector. The kinetic parameters of the RCANA may be varied in response to the amount of an effector , which may be an allosteric effector molecule. Just as allosteric protein enzymes undergo a change in their kinetic parameters or of their enzymatic activity in response to interactions with an effector molecule, the catalytic abilities of RCANAs may be similarly modulated by effectors. As demonstrated herein, the effectors may be small molecules, proteins, peptides or molecules that interact with proteins, peptides or other molecules. RCANAs transduce molecular recognition into catalysis upon interaction with an effector that interacts with a portion of the RCANA.

[0095] As used herein, the term "effector," "effector molecule", "allosteric effector" or "allosteric effector molecule" means a molecule or process that changes the kinetic parameters or catalytic activity of an RCANA.

[0096] As used herein, the term "catalytic residue" refers to residues that when mutated decrease the activity of the RCANA. Mutating a residue that affects the catalytic activity of a ribozyme following the selection of the RCANA, may cause different residues to become sensitive to mutation than in the original ribozyme. The relative mutational sensitivity of a given 'catalytic residue' may change before and after the selection of the RCANA. These secondary mutations are also encompassed by the present invention.

[0097] As used herein, the term "aptamer" refers to a nucleic acid that has been specifically selected to optimally bind to a target ligand. As described above, it is important to recognize that an aptamer is fundamentally different than an RCANA.

[0098] As used herein, the term "kinetic parameters" refers to any aspect of the catalytic activity of the nucleic acid. Changes in the kinetic parameters of a catalytic RCANA produce changes in the catalytic activity of the RCANA such as a change in the rate of reaction or a change in substrate specificity. For example, self-splicing of an RCANA out of a gene environment may result from a change in the kinetic parameters of the RCANA.

[0099] As used herein, the term "catalytic" or "catalytic activity" refers to the ability of a substance to affect a change in itself or of a substrate under permissive conditions. As used herein, the term "protein-protein complex" or "protein complex" refers to an association of more than one protein. The proteins that make up a protein complex may be associated by functional, stereochemical, conformational, biochemical, or electrostatic mechanisms. It is intended that the term encompass associations of any number of proteins.

[00100] As used herein, the term "*in vivo*" refers to cellular systems and organisms, e.g., cultured cells, yeast, bacteria, plants and/or animals.

[00101] As used herein the terms "protein", "polypeptide" or "peptide" refer to compounds comprising amino acids joined via peptide bonds and are used interchangeably. As used herein, the term "endogenous" refers to a substance the source of which is from within a cell, cell extract or reaction system. Endogenous substances are produced by the metabolic activity of, e.g., a cell. Endogenous substances, however, may nevertheless be produced as a result of manipulation of cellular metabolism to, for example, make the cell express the gene encoding the substance.

[00102] As used herein, the term "exogenous" refers to a substance the source of which is external to a cell, cell extract or reaction system. An exogenous substance may nevertheless be internalized by a cell by any one of a variety of metabolic or induced means known to those skilled in the art.

[00103] As used herein the term "modified base" refers to a non-natural nucleotide of any sort, in which a chemical modification may be found on the nucleobase, the sugar, or the polynucleotide backbone or phosphodiester linkage.

[00104] As used herein, the term "gene" means the coding region of a deoxyribonucleotide sequence encoding the amino acid sequence of a protein. The term includes sequences located adjacent to the coding region on both the 5, and 3, ends such that the deoxyribonucleotide sequence corresponds to the length of the full-length mRNA for the protein. The term "gene" encompasses both cDNA and genomic forms of a gene. A genomic form or clone of a gene contains the coding region interrupted with non-coding

sequences termed "introns" or "intervening regions" or "intervening sequences." Introns are segments of a gene that are transcribed into nuclear RNA (hnRNA); introns may contain regulatory elements such as enhancers. Introns are removed, excised or "spliced out" from the nuclear or primary transcript; introns therefore are absent in the messenger RNA (mRNA) transcript. The mRNA functions during translation to specify the sequence or order of amino acids in a nascent polypeptide.

[00105] In addition to containing introns, genomic forms of a gene may also include sequences located on both the 5' and 3' end of the sequences that are present on the RNA transcript. These sequences are referred to as "flanking" sequences or regions (these flanking sequences are located 5' or 3' to the non-translated sequences present on the MRNA transcript). The 5' flanking region may contain regulatory sequences such as promoters and enhancers that control or influence the transcription of the gene. The 3' flanking region may contain sequences that direct the termination of transcription, post-transcriptional cleavage and polyadenylation.

[00106] DNA molecules are said to have "5' ends" and "3' ends" because mononucleotides are reacted to make oligonucleotides in a manner such that the 5' phosphate of one mononucleotide pentose ring is attached to the 3' oxygen of its neighbor in one direction via a phosphodiester linkage. Therefore, an end of an oligonucleotides referred to as the "5' end" if its 5' phosphate is not linked to the 3' oxygen of a mononucleotide pentose ring and as the "3' end" if its 3' oxygen is not linked to a 5' phosphate of a subsequent mononucleotide pentose ring. As used herein, a nucleic acid sequence, even if internal to a larger oligonucleotide, also may be said to have 5' and 3' ends. In either a linear or circular DNA molecule, discrete elements are referred to as being "upstream" or 5' of the "downstream" or 3' elements. This terminology reflects the fact that transcription proceeds in a 5' to 3' fashion along the DNA strand.

[00107] The term "gene of interest" as used here refers to a gene, the function and/or expression of which is desired to be investigated, or the expression of which is desired to be regulated, by the present invention. In the present disclosure, the td gene of the T4 bacteriophage is an example of a gene of interest and is described herein to illustrate the

invention. The present invention may be useful in regard to any gene of any organism, whether of a prokaryotic or eukaryotic organism.

[00108] The term "hybridize" as used herein, refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing. Hybridization and the strength of hybridization (i.e., the strength of the association between the nucleic acid strands) is impacted by such factors as the degree of complementary between the nucleic acids, stringency of the conditions involved, the melting temperature of the formed hybrid, and the G:C (or U:C for RNA) ratio within the nucleic acids.

[00109] The terms "complementary" or "complementarity" as used herein, refer to the natural binding of polynucleotides under permissive salt and temperature conditions by base-pairing. For example, for the sequence "A-G-T" binds to the complementary sequence "T-C-A". Complementarity between two single-stranded molecules may be partial, in which only some of the nucleic acids bind, or it may be complete when total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of hybridization between nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acids strands.

[00110] The term "homology," as used herein, refers to a degree of complementarity. There may be partial homology or complete homology (i.e., identity). A partially complementary sequence is one that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid; it is referred to using the functional term "substantially homologous." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or Northern blot, solution hybridization and the like) under conditions of low stringency. A substantially homologous sequence or probe will compete for and inhibit the binding (i.e., the hybridization) of a completely homologous sequence or probe to the target sequence under conditions of low stringency. This is not to say that conditions of low stringency are such that non-specific binding is permitted; low stringency conditions require that the binding of two sequences to one another be a specific (i.e., selective) interaction. The absence of non-

specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% identity); in the absence of non-specific binding, the probe will not hybridize to the second non-complementary target sequence. When used in reference to a single-stranded nucleic acid sequence, the term "substantially homologous" refers to any probe which can hybridize (i.e., it is the complement of) the single-stranded nucleic acid sequence under conditions of low stringency as described.

[00111] As known in the art, numerous equivalent conditions may be employed to comprise either low or high stringency conditions. Factors such as the length and nature (DNA, RNA, base composition) of the sequence, nature of the target (DNA, RNA, base composition, presence in solution or immobilization, etc.), and the concentration of the salts and other components (e.g., the presence or absence of formamide, dextran sulfate and/or polyethylene glycol) are considered and the hybridization solution may be varied to generate conditions of either low or high stringency different from, but equivalent to, the above listed conditions.

[00112] As used herein the term "stringency" is used in reference to the conditions of temperature, ionic strength, and the presence of other compounds such as organic solvents, under which nucleic acid selections are conducted. With "high stringency" conditions a relatively small number of nucleic acid catalysts will be selected from a random sequence pool, while under "low stringency conditions a larger number of nucleic acid catalysts will be selected from a random sequence pool.

[00113] Numerous equivalent conditions may be employed to comprise low or high stringency conditions; factors such as the length of incubation of the reaction, the presence of competitive inhibitors of the reaction, the buffer conditions under which the reaction is carried out, the temperature at which the reaction is carried out are considered and the hybridization solution may be varied to generate conditions of low stringency selection different from, but equivalent to, the above listed conditions.

[00114] The term "antisense," as used herein, refers to nucleotide sequences that are complementary to a specific DNA or RNA sequence. The term "antisense strand" is used in

reference to a nucleic acid strand that is complementary to tile "sense" strand. Antisense molecules may be produced by any method, including synthesis by ligating the gene(s) of interest in a reverse orientation to a viral promoter that permits the synthesis of a complementary strand. Once introduced into a cell, the transcribed strand combines with natural sequences produced by the cell to form duplexes. These duplexes then block either the further transcription or translation. In this manner, mutant phenotypes may also be generated. The designation "negative" is sometimes used in reference to the antisense strand, and "positive" is sometimes used in reference to the sense strand. The term is also used in reference to RNA sequences that are complementary to a specific RNA sequence (e.g., mRNA). Included within this definition are antisense RNA ("asRNA") molecules involved in genetic regulation by bacteria.

[00115] Antisense RNA may be produced by any method, including synthesis by splicing the gene(s) of interest in a reverse orientation to a viral promoter that permits the synthesis of a coding strand. Once introduced into an embryo, this transcribed strand combines with natural mRNA produced by the embryo to form duplexes. These duplexes then block either the further transcription of the mRNA or its translation. In this manner, mutant phenotypes may be generated. The term "antisense strand" is used in reference to a nucleic acid strand that is complementary to the "sense" strand. The designation. (-) (i.e., "negative") is sometimes used in reference to the antisense strand with the designation (+) sometimes used in reference to the sense (i.e., "positive") strand.

[00116] A gene may produce multiple RNA species that are generated by differential splicing of the primary RNA transcript. cDNAs that are splice variants of the same gene will contain regions of sequence identity or complete homology (representing the presence of the same exon or portion of the same exon on both cDNAs) and regions of complete non-identity (for example, representing the presence of exon "A" on cDNA 1 wherein cDNA 2 contains exon "B" instead). Because the two cDNAs contain regions of sequence identity they will both hybridize to a probe derived from the entire gene or portions of the gene containing sequences found on both cDNAs; the two splice variants are therefore substantially homologous to such a probe and to each other.

[00117] "Transformation," as defined herein, describes a process by which exogenous DNA enters and changes a recipient cell. It may occur under natural or artificial conditions using various methods well known in the art. Transformation may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method is selected based on the host cell being transformed and may include, but is not limited to, viral infection, electroporation, lipofection, and particle bombardment. Such "transformed" cells include stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome. The term "transfection" as used herein refers to the introduction of foreign DNA into eukaryotic cells.

[00118] Transfection may be accomplished by a variety of methods known to the art including, e.g., calcium phosphate-DNA co-precipitation, DEAE-dextran-mediated transfection, polybrene-mediated transfection, electroporation, microinjection, liposome fusion, lipofection, protoplast fusion, retroviral infection, and biolistics. Thus, the term "stable transfection" or "stably transfected" refers to the introduction and integration of foreign DNA into the genome of the transfected cell. The term "stable transfectant" refers to a cell that has stably integrated foreign DNA into the genomic DNA. The term also encompasses cells that transiently express the inserted DNA or RNA for limited periods of time. Thus, the term "transient transfection" or "transiently transfected" refers to the introduction of foreign DNA into a cell where the foreign DNA fails to integrate into the genome of the transfected cell. The foreign DNA persists in the nucleus of the transfected cell for several days. During this time the foreign DNA is subject to the regulatory controls that govern the expression of endogenous genes in the chromosomes. The term "transient transfectant" refers to cells that have taken up foreign DNA but have failed to integrate this DNA.

[00119] As used herein, the term "selectable marker" refers to the use of a gene that encodes an enzymatic activity and which confers the ability to grow in medium lacking what would otherwise be an essential nutrient (e.g., the HIS3 gene in yeast cells); in addition, a selectable marker may confer resistance to an antibiotic or drug upon the cell in which the selectable marker is expressed. A review of the use of selectable markers in mammalian cell

lines is provided in Sambrook, J. et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York (1989) pp.16.9-16.15.

[00120] As used herein, the term "reporter gene" refers to a gene that is expressed in a cell upon satisfaction of one or more contingencies and which, upon expression, confers a detectable phenotype to the cell to indicate that the contingencies for expression have been satisfied. For example, the gene for Luciferase confers a luminescent phenotype to a cell when the gene is expressed inside the cell. In the present invention, the gene for Luciferase may be used as a reporter gene such that the gene is only expressed upon the splicing out of an intron in response to an effector. Those cells in which the effector activates splicing of the intron will express Luciferase and will glow.

[00121] As used herein, the term "vector" is used in reference to nucleic acid molecules that transfer DNA segment(s) from one cell to another. The term "vehicle" is sometimes used interchangeably with "vector." The term "vector" as used herein also includes expression vectors in reference to a recombinant DNA molecule containing a desired coding sequence and appropriate nucleic acid sequences necessary for the expression of the operably linked coding sequence in a particular host organism. Nucleic acid sequences necessary for expression in prokaryotes usually include a promoter, an operator (optional), and a ribosome binding site, often along with other sequences. Eukaryotic cells are known to utilize promoters, enhancers, and termination and polyadenylation signals.

[00122] As used herein, the term "amplify", when used in reference to nucleic acids refers to the production of a large number of copies of a nucleic acid sequence by any method known in the art. Amplification is a special case of nucleic acid replication involving template specificity. Template specificity is frequently described in terms of "target" specificity. Target sequences are "targets" in the sense that they are to be sorted out from other nucleic acid. Amplification techniques have been designed primarily for this sorting out.

[00123] As used herein, the term "primer" refers to an oligonucleotide, whether occurring naturally as in a purified restriction digest or produced synthetically, which is capable of acting as a point of initiation of synthesis when placed under conditions in which

synthesis of a primer extension product which is complementary to a nucleic acid strand is induced, (i.e., in the presence of nucleotides and an inducing agent such as DNA polymerase and at a suitable temperature and pH). The primer may be single stranded for maximum efficiency in amplification but may alternatively be double stranded. If double stranded, the primer is first treated to separate its strands before being used to prepare extension products. The primer must be sufficiently long to prime the synthesis of extension products in the presence of the inducing agent. The exact lengths of the primers will depend on many factors, including temperature, source of primer and the use of the method.

[00124] As used herein, the term "probe" refers to an oligonucleotide (i.e., a sequence of nucleotides), whether occurring naturally as in a purified restriction digest or produced synthetically, recombinantly or by PCR amplification, which is capable of hybridizing to another oligonucleotide of interest. A probe may be single-stranded or double-stranded. Probes are useful in the detection, identification and isolation of particular gene sequences. It is contemplated that any probe used in the present invention will be labeled with any "reporter molecule," so that is detectable in any detection system, including, but not limited to enzyme (e.g. ELISA, as well as enzyme-based histochemical assays), fluorescent, radioactive, and luminescent systems. It is not intended that the present invention be limited to any particular detection system or label.

[00125] As used herein, the term "target" when used in reference to the polymerase chain reaction, refers to the region of nucleic acid bounded by the primers used for polymerase chain reaction. Thus, the "target" is sought to be sorted oat from other nucleic acid sequences. A "segment" is defined as a region of nucleic acid within the target sequence.

[00126] As used herein, the term "polymerase chain reaction" ("PCR") refers to the method of K.B. Mullis U.S. Patent Nos. 4,683,195, 4,683,202, and 4,965,188, hereby incorporated by reference, which describe a method for increasing the concentration of a segment of a target sequence in a mixture of genomic DNA without cloning or purification. This process for amplifying the target sequence consists of introducing a large excess of two oligonucleotide primers to the DNA mixture containing the desired target sequence, followed

by a precise sequence of thermal cycling in the presence of a DNA polymerase. The two primers are complementary to their respective strands of the double stranded target sequence.

[00127] To effect amplification, the mixture is denatured and the primers then annealed to their complementary sequences within the target molecule. Following annealing, the primers are extended with a polymerase so as to form a new pair of complementary strands. The steps of denaturation, primer annealing and polymerase extension can be repeated many times (i.e., denaturation, annealing and extension constitute one "cycle"; there can be numerous "cycles") to obtain a high concentration of an amplified segment of the desired target sequence. The length of the amplified segment of the desired target sequence is determined by the relative positions of the primers with respect to each other, and therefore, this length is a controllable parameter. By virtue of the repeating aspect of the process, the method is referred to as the "polymerase chain reaction" (hereinafter "PCR"). Because the desired amplified segments of the target sequence become the predominant sequences (in terms of concentration) in the mixture, they are said to be "PCR amplified".

[00128] With PCR, it is possible to amplify a single copy of a specific target sequence in genomic DNA to a level detectable by several different methodologies (e.g., hybridization with a labeled probe; incorporation of biotinylated primers followed by avidin-enzyme conjugate detection; incorporation of ^{32}P-labeled deoxynucleotide triphosphates, such as DCTP or DATP, into the amplified segment). In addition to genomic DNA, any oligonucleotide sequence can be amplified with the appropriate set of primer molecules. In particular the amplified segments created by the PCR process itself are, themselves, efficient templates for subsequent PCR amplifications.

EXAMPLE 1: GPITH1P6

ENGINEERING OF AN RCANA FOR *IN VIVO* DETECTION APPLICATIONS

[00129] The first example illustrates how to make an RCANA construct and demonstrates self-splicing of the RCANA out of a gene in response to an effector molecule.

[00130] Construction of a RCANA. Oligos GpIWt3.129: 5'-TAA TCT TAC CCC GGA ATT ATA TCC AGC TGC ATG TCA CCA TGC AGA GCA GAC TAT ATC TCC AAC TTG TTA AAG CAA GTT GTC TAT CGT TTC GAG TCA CTT GAC CCT ACT CCC CAA AGG GAT AGT CGT TAG-3' (SEQ ID NO: 1) and

[00131] GpITh1P6.131: 5'-GCC TGA GTA TAA GGT GAC TTA TAC TTG TAA TCT ATC TAA ACG GGG AAC CTC TCT AGT AGA CAA TCC CGT GCT AAA TTA TAC CAG CAT CGT CTT GAT GCC CTT GGC AGA TAA ATG CCT AAC GAC TAT CCC TT-3' (SEQ ID NO: 2) were annealed and extended in a 30 µl reaction containing 100 pmoles of each oligo, 250 mM Tris-HCl (pH 8.3), 40 mM $MgCl_2$, 250 mM NaCl, 5 mM DTT, 0.2 mM each dNTP, 45 units of AMV reverse transcriptase (RT: Amersham Pharmacia Biotech, Inc., Piscataway, NJ) at 37° C for 30 minutes. The extension reaction was diluted 1 to 50 in H_2O.

[00132] A PCR reaction containing 1 µl of the extension dilution, 500 mM KCl, 100 mM Tris-HCl, (pH 9.0), 1% Triton® x-100, 15 mM $MgCl_2$, 0.4 µM of GpIWt1.75: 5'-GAT AAT ACG ACT CAC TAT AGG GAT CAA CGC TCA GTA GAT GTT TTC TTG GGT TAA TTG AGG CCT GAG TAT AAG GTG-3' (SEQ ID NO:3), 0.4 µM of GpIWt4.89: 5'-CTT AGC TAC AAT ATG AAC TAA CGT AGC ATA TGA CGC AAT ATT AAA CGG TAG CAT TAT GTT CAG ATA AGG TCG TTA ATC TTA CCC CGG AA-3' (SEQ ID NO:4), 0.2 mM each dNTP and 1.5 units of Taq polymerase (Promega, Madison, WI) was thermocycled 20 times under the following regime: 94° C for 30 seconds, 45° C for 30 seconds, 72° C for 1 minute. The PCR reaction was precipitated in the presence of 0.2 M NaCl and 2.5 volumes of ethanol and then quantitated by comparison with a molecular weight standard using agarose gel electrophoresis.

[00133] The RCANA construct was transcribed in a 10 µl high yield transcription reaction (AmpliScribe from Epicentre, Madison, WI. The reaction contained 500 ng PCR product, 3.3 pmoles of ^{32}P [^{32}P]UTP, 1X AmpliScribe transcription buffer, 10 mM DTT, 7.5 mM each NTP, and 1 µl AmpliScribe T7 polymerase mix. The transcription reaction was incubated at 37° C for 2 hours. One unit of RNase free-DNase was added and the reaction returned to 37° C for 30 minutes. The transcription was then purified on a 6% denaturing

polyacrylamide gel to separate the full length RNA from incomplete transcripts and spliced products, eluted and quantitated spectrophotometrically.

[00134] *In vitro* Assay. The RNA (4 pmoles/12 µl H_2O) was heated to 94° C for 1 minute then cooled to 37° C over 2 minutes in a thermocycler. The RNA was divided into 2 splicing reactions (9 µl each) containing 100 mM Tris-HCl (pH 7.45), 500 mM KCl and 15 mM $MgCl_2$, plus or minus theophylline (2 mM). The reactions were immediately placed on ice for 30 minutes. GTP (1 mM) was added to the reactions (final volume of 10 µl) and the reactions were incubated at 37° C for 2 hours.

[00135] The reactions were terminated by the addition of stop dye (10 µl) (95% formamide, 20 mM EDTA, 0.5% xylene cyanol, and 0.5% bromophenol blue). The reactions were heated to 70° C for 3 minutes and 10 µl was electrophoresed on a 6% denaturing polyacrylamide gel. The gel was dried, exposed to a phosphor screen and analyzed using a Molecular Dynamics Phosphorimager (Sunnyvale, CA).

[00136] Activation was determined from the amount of circular intron in each reaction. Circularized introns migrate slower than linear RNA and can be seen as the bands above the dark bands of linear RNA in the +Theo lanes of the gels of Figures 2a and. 2b.

[00137] *In vivo* Screening of Group I Aptazymes. The RCANA constructs as well as the wild type and a negative control were ligated into a vector that contains the T4 td intron with Eco RI and Spe I flanking the P6 region, transformed and minipreped. The plasmids were then transformed into C600:Thy A Kan^R cells (cells lacking thymidine synthetase). Individual colonies were picked and grown in rich media overnight. Theophylline (1 µl: 6.6 mM) or H_2O (1 µl) was added to 2 µl of the overnight growth and was spotted on either minimal media plates, or minimal media plates with thymine, see Figure 3.

EXAMPLE 2: GPIP6THPOOL

IN VITRO SELECTION TO OPTIMIZE AN RCANA FOR *IN VIVO* DETECTION APPLICATIONS

[00138] Example 2 illustrates how to generate a population of RCANA so that there is variation in the nucleotide sequence of the aptamers. This example also illustrates how to select for phenotypes that are responsive to an effector molecule from among that population of RCANA.

[00139] Construction of Pool. The construction of the pool was similar to the construction of the individual engineered RCANA constructs. Oligos GpIWt3.129 and GpIThP6pool: 5'-GCC TGA GTA TAA GGT GAC TTA TAC TAG TAA TCT ATC TAA ACG GGG AAC CTC TCT AGT AGA CAA TCC CGT GCT AAA TN(1-4)A TAC CAG CAT CGT CTT GAT GCC CTT GGC AGN(1-4) TAA ATG CCT AAC GAC TAT CCC TT-3' (SEQ ID NO:5) were extended in the same manner as above. The extension reaction was diluted and used as template for a PCR reaction. The PCR reaction was similar to the reaction described with the following exceptions: the volume was doubled and GpIWt4.89 was replaced with Gp1MutG.101: 5'-CTT AGC TAC AAT ATG AAC TAA CGT AGC ATA TGA CGC AAT ATT AAA CGG TAG TAT TAT GTT CAG ATA AGG TCG TTA ATC TTA CCC CGG AAT TCT ATC CAG CT-3' (SEQ ID NO:6) in which there is an G to A mutation at the terminal residue of the intron. The pool had a diversity of $1.16X10^5$ molecules. RNA was made as described above.

[00140] *In vitro* Negative Selection. The RNA (10 pmoles/70 µl H_2O) was heated to 94° C for 1 minute then cooled to 37° C over 2 minutes in a thermocycler. The splicing reaction (90 µl) contained 100 mM Tris-HCl (pH 7.45), 500 mM KCl and 15 mM $MgCl_2$. The reaction was immediately placed on ice for 30 minutes. GTP (1 mM) was added to the reaction (final volume of 100 µl) and the reaction was incubated at 37° C for 20 hours. The reaction was terminated by the addition 20 mM EDTA and precipitated in the presence of 0.2 M NaCl and 2.5 volumes of ethanol. The reaction was resuspended in 10 µl H_2O and 10 µl

stop dye and heated to 70° C for 3 minutes and was electrophoresed on a 6% denaturing polyacrylamide gel with Century™Marker ladder (Ambion, Austin, TX). The gel was exposed to a phosphor screen and analyzed. The unreacted RNA was isolated from the gel, precipitated and resuspended in 10 µl of H_2O.

[00141] Positive Selection. The RNA (5 µl of negative selection) was heated to 94° C for 1 minute then cooled to 37° C over 2 minutes in a thermocycler. The positive splicing reaction (45 µl) contained 100 mM Tris-HCl (pH 7.45), 500 mM KCl, 15 mM $MgCl_2$ and 1mM theophylline. The reaction was immediately placed on ice for 30 minutes. GTP (1 mM) was added to the reaction (final volume of 50 µl) and the reaction was incubated at 37° C for 1 hour. The reaction was terminated by the addition of stop dye, heated to 70° C for 3 minutes and was electrophoresed on a 6% denaturing polyacrylamide gel with Century™Marker ladder. The gel was exposed to a phosphor screen and analyzed. The band corresponding to the linear intron was isolated from the gel and precipitated and resuspended in 20 µl H_2O.

[00142] Amplification and Transcription. The RNA was reverse transcribed in a reaction containing 250 mM Tris-HCl (pH 8.3), 375 mM KCl, 15 mM, $MgCl_2$, 0.1 M DTT, 0.4 mM of each dNTP 2 µM GpIMutG.101 and 400 units of SuperScript II reverse transcriptase (Gibco BRL, Rockville, MD). The cDNA was then PCR amplified, transcribed and gel purified as described above.

[00143] Figure 3 depicts an *in vivo* assay system for Group I introns of the present invention. The td intron normally sits within the td gene for thymidylate synthase (TS) in phage T4. A ThyA E. coli host that lacks cellular TS is unable to grow in the absence of exogenous thymine or thymidine (-Thy). The cloned td gene can complement the ThyA cells and grow on -Thy media. Conversely, cells that lack TS have a selective advantage on media containing thymidine and trimethoprim. Therefore, cells harboring theophylline-responsive Group I aptazymes grow better in the presence of theophylline and the absence of thymidine. In contrast, the same cells grow better in the absence of theophylline and the presence of thymidine and trimethoprim.

[00144] This strategy provides both a positive *in vivo* screen and selection for theophylline-dependent activation and a negative *in vivo* screen and selection for theophylline-absent repression. The assay system of Figure 3 was used in Example 1, above, for the *in vivo* screening of Group I aptazymes in a specific embodiment of the present invention.

[00145] Figure 4a depicts the critical residues of the P6 region of the Group I ribozyme joined to the anti-theophylline aptamer by a short randomized region to generate a pool of RCANA of the present invention. The residues shown in bold in Figure 4a are the P6 critical residues, and the faded residues shown in Figure 4a are the anti-theophylline aptamer. The randomized regions are designated in Figure 4a as "N1-4". Approximately 40 random sequence residues are introduced into the N1-4 region of the construct to synthesize a pool of RCANA, referred to herein as a communication module pool.

EXAMPLE 3

POLYPEPTIDE DEPENDENT REGULATABLE, CATALYTICALLY ACTIVE NUCLEIC ACIDS

[00146] Natural nucleic acids frequently rely on proteins for stabilization or catalytic activity. In contrast, nucleic acids selected *in vitro* can catalyze a wide range of reactions even in the absence of proteins. In order to augment selected nucleic acids with protein functionalities, the present invention includes a technique for the selection of protein-dependent ribozyme ligases.

[00147] The catalytic domain of the ribozyme ligase, L1, was randomized, and variants that required one of two protein cofactors, a tyrosyl tRNA synthetase (Cyt18) or hen egg white lysozyme, were selected. The resultant regulatable, catalytically active nucleic acids were activated thousands of fold by their cognate, protein effectors, and could specifically recognize the structures of the native proteins. Protein-dependent regulatable, catalytically active nucleic acids are adaptable to novel assays for the detection of target proteins, and the generality of the selection method, as demonstrated herein allows for the

identification of regulatable, catalytically active nucleic acids using high-throughput methods and equipment. These regulatable, catalytically active nucleic acids are able to, for example, recognize a sizable fraction of a proteome.

[00148] It has been recognized that it is possible to design and select effector-modulated ribozymes (RCANA) that show astounding activation parameters relative to allosteric proteins. For example, the inventors recognized that Breaker and his co-workers engineered an allosteric hammerhead ribozyme that is inhibited by 180-fold in the presence of a small molecule, ATP (Tang, J. & Breaker, R.R. Rational design of allosteric ribozymes. *Chem. Biol.* 4, 453-459 (1997)). The present inventors had also engineered an effector-activated ribozyme ligase that is activated by 1,600-fold in the presence of theophylline (Robertson, M.P. & Ellington, A.D., Design and optimization of effector-activated ribozyme ligases. *Nucleic Acids Res.* 28, 1751-1759 (2000)). Allosteric domains have also been selected from random sequence pools appended to the hammerhead ribozyme; these domains mediate a 5,000-fold activation of the ribozyme by other small molecules, e.g., cyclic nucleotide monophosphates (Koizumi, M., Soukup, G.A., Kerr, J.N. & Breaker, R.R., Allosteric selection of ribozymes that respond to the second messengers cGMP and cAMP. *Nat. Struct. Biol.* 6, 1062-1071 (1999)).

[00149] The present inventors recognized and herein demonstrate that it is possible to identify not only ribozymes, but nucleic acid segments that are activated by protein effectors. They further recognized that previous attempts to isolate ribozymes had required active catalytic domains within those ribozymes. All previously isolated ribozymes had been designed, modified, isolated or identified with natural or enhanced catalytic domains, hence the isolation of these ribozymes are extremely dependent on the catalytic domain for their isolation.

[00150] The RNAse P ribozyme from eubacteria has been shown to catalyze the cleavage of tRNA, it is normally complexed with a protein (P-protein) that substantially enhances its activity. Similarly, the Group I intron ND1 is extremely dependent on Cyt18, a tyrosyl tRNA synthetase from *Neurospora crassa* mitochondria, while the tertiary structure of the intron bI5 is stabilized by its cognate protein, CBP2. Proteins have been frequently

found to assist in the folding of RNA molecules, acting as chaperons to partially solvate the polyanionic backbone (Weeks, K.M. Protein-facilitated RNA folding. *Curr. Opin. Struct. Biol.* 7, 336-342 (1997)).

[00151] The present invention includes a generalized selection scheme for the isolation of regulatable, catalytically active nucleic acids. Using the present invention a novel class of not just ribozymes, but rather, regulatable, catalytically active nucleic acids that are specifically activated thousands of fold by protein effectors such as Cyt18 and lysozyme have been create, isolated and identified.

[00152] *In vitro* selection of protein-dependent ribozymes. While attempting to identify peptide- and protein-dependent ribozymes the present inventors used novel strategies for the design and selection of ribozymes that were activated by small molecular effectors. However, when peptide- and protein-binding sites were appended to stem C of the small L1 ligase (Figure 17A) little or no modulation of activity was observed in the presence of cognate peptide or protein effectors (data not shown). Similarly, when a random sequence loop was introduced at the termini of stem C, selection for protein-dependent variants produced only very modest activation (< 2x).

[00153] It was then discovered that engineering protein-dependent ribozymes required fundamentally different principles than engineering small molecule-dependent ribozymes. In particular, it was recognized that small molecules that bind to limited allosteric sites in turn to potentiate small but significant reorganizations of the secondary and tertiary structures of core ribozymes. It was further discovered that larger effector molecules, such as proteins, bind to much larger sites and might sterically inhibit the catalytic core. Therefore, it was necessary to include the catalytic core in the selection. To this end, a nucleic acid segment pool based on the L1 ligase (L1-N50) in which critical catalytic residues were also randomized (Figure 17B) was designed.

[00154] The L1-N50 pool (10^{15} starting species) was subjected to an iterative regime of negative and positive selections for ligation activity (Figure 17C). The pool was initially incubated with a biotinylated substrate and reactive species were removed; the pool was then mixed with the effector molecule, a tyrosyl tRNA synthetase from *Neurospora* mitochondria

(Cyt18), and reactive species were removed and amplified. The Cyt18 protein was chosen as an effector because it was known to both tightly bind (Kd in the femptomolar range) and activate a natural RNA catalyst, a group I self-splicing intron. During the course of these studies, and in negative selection screens in general using the present invention, the stringency of the negative selections may be increased by increasing the time allowed for ligation and substrate concentration in the absence of Cyt18. Conversely, the stringency of the positive selections may increased by decreasing the time allowed for ligation and the substrate concentration (Figure 18A).

[00155] The degree of protein-dependent activation was assessed in a standard assay, and progressively increased from Round 5 onwards (Figure 18B). By Round 7, protein-dependent activation was greater than 50,000-fold. At the conclusion of the selection it had risen to over 75,000-fold. The most prevalent clone in the selected population (cyt7-2) performed the ligation reaction with an observed rate of 1.6 h^{-1} in the presence of Cyt18, but this rate dropped to 0.00005 h^{-1} when the protein was left out of the reaction, a difference of 34,000-fold. Another clone (cyt9-18) from the selection had even better activation parameters, ligating at a rate of 2.1 h^{-1} with Cyt18 included in the reaction, but only 0.00002 h^{-1} without protein for a difference of 94,000-fold. Importantly, these values are many orders of magnitude greater than the known ligand-mediated activation of allosteric protein enzymes, and are 10- to 100-fold greater than the previously observed activation of ribozymes by small molecule effectors.

[00156] While the extent of Cyt18 activation of the aptazyme ligase was impressive, Cyt18 had previously been shown to similarly activate a group I self-splicing intron. In order to determine whether the ability to select for protein-dependent activation of ribozyme catalysis was specific to certain types of proteins or was a more general phenomena, ribozyme ligases that could be activated by a protein not normally known to bind RNA, hen egg white lysozyme were isolated. Using the same selection scheme and progressive increases in stringency (Figure 18C), regulatable, catalytically active nucleic acids that were activated by lysozyme were isolated in 11 cycles of selection and amplification. The final, selected population was activated about 800-fold by lysozyme (Figure 18D) and an isolated

clone, lys11-2, exhibited a 3100-fold activation, ligating with an observed rate of 0.6 h^{-1} in the presence of lysozyme but only 0.0002 h^{-1} without lysozyme.

[00157] Characterization of protein-dependent ribozymes. Individual ribozymes were cloned from both selections and sequenced (Figure 19A). In both instances, only a few families of ribozymes remained. These results are more in line with those previously observed for ribozyme selections with small organic ligands. Using the present invention, individual sequences could be folded to fit within the general structural context of the L1 ligase (Figure 19B). The selected ribozymes were still highly dependent on the presence of the 3' primer for activity, as was the parental L1 ligase. The selected sequences were hypothesized to form extended 'stem C' structures. The formation of such extended stems was again consistent with L1 ligase.

[00158] The distal portion of stem C, adjacent to the hairpin, was not conserved following partial randomization and re-selection, indicating that this portion of the ribozyme was not critical for activity. Moreover, the distal, hairpin portion of stem C can be shortened without loss of activity, and the hairpin may be replaced by aptamers that bind small organic ligands to generate regulatable, catalytically active nucleic acids. While the internal loop region of stem C, adjacent to the 3-arm junction, was conserved following doped sequence selection, complete randomization of this region followed by selection for ligase function yielded a variety of sequence solutions. Therefore, the selected protein-dependent ribozymes differed substantially from the parental ribozyme in this region.

[00159] Specificity of activation. In order to assess the specificity of activation of selected ribozymes by protein effectors, the Cyt18-dependent population was incubated with a variety of proteins, including lysozyme, *E. coli* tryptophanyl tRNA synthetase, ricin A chain, and MS2 coat protein. No activation was observed with proteins that were not used during the isolation. Similarly, lysozyme-dependent clones were incubated with Cyt18, turkey lysozyme, and lysozyme from human milk. Only the extremely homologous (98%) turkey lysozyme showed cross-activation, while the other protein effectors were inactive. Therefore, activation is highly specific, and activation by some contaminating factor (salt, magnesium) that might have been introduced during protein preparations is unlikely. In

addition, as several of the non-cognate proteins were known to bind RNA both specifically and non-specifically, general stabilization of ribozyme structure by protein 'salts' is also an unlikely mechanism for activation.

[00160] Nonetheless, it was still possible that contaminants unique to each protein preparation were responsible for activation. In order to discount this source for cross-reactivity, the regulatable, catalytically active nucleic acids were incubated with inactivated cognate proteins (data not shown). Cyt18 was denatured either by heating or by incubation with sodium dodecly sulfate (SDS), while lysozyme was denatured by a combination of disulfide bond reduction and heating. Denatured Cyt18 was unable to activate ribozyme catalysis, while only lysozyme that had been both reduced and denatured was unable to activate catalysis. Both reduction and denaturation are required to eliminate lysozyme activity. It appeared as though the selected ribozymes were not only specific for their protein effectors, but may also be dependent on protein conformation. In fact, given that anti-peptide antibodies have been shown to partially denature protein structure it may be that protein-activated ribozymes will be found to be even more sensitive to protein conformation than other proteins.

[00161] Next, the inventors probed the activation of individual regulatable, catalytically active nucleic acids by using RNA inhibitors of the protein effectors. Previously selected both anti-Cyt18 (data not shown) and anti-lysozyme aptamers were used under buffer conditions similar to those used for these selections. These and other RNA molecules were incubated together with regulatable, catalytically active nucleic acids and their protein effectors, and protein-dependent activation was assessed. Several RNA molecules slightly reduced Cyt18 activation of clone cyt7-2, possibly due to non-specific competition for binding. However, the greatest reduction in activity was observed with RNAs known to bind specifically to Cyt18. The ND1 intron is an *in vivo* substrate for Cyt18 and shows the greatest reduction in activity, while an aptamer that has been shown to inhibit the ability of Cyt18 to interact with ND1 (M12; Cox and Ellington, unpublished results) was also an effective inhibitor. In contrast, an aptamer that binds to Cyt18 but does not inhibit its interactions with ND1 (B17; data not shown) inhibits activation no better than: an anti-lysozyme aptamer (c1), a random sequence pool (N30), or tRNA. Lysozyme activation of its

corresponding regulatable, catalytically active nucleic acids (lys11-2) proved to be relatively impervious to all inhibitors except for a high affinity anti-lysozyme aptamer (c1, K_d = 31 nM), which reduced activation to background levels. The specificity of inhibition observed with these different RNA species further emphasizes the specificity of the interactions between effector proteins and their cognate regulatable, catalytically active nucleic acids.

[00162] A direct correlation between the lysozyme binding and ribozyme activation could be demonstrated (Figure 21). Lysozyme interacts with its regulatable, catalytically active nucleic acids with an apparent K_d of 1.5 μM, while the Cyt18 regulatable, catalytically active nucleic acids could not be saturated even at protein concentrations up to 2.5 μM). Moreover, when the activity of a lysozyme-dependent ribozyme was assayed as a function of salt concentration, binding and catalysis were both depressed by high (1 M) salt concentrations (data not shown). Interestingly, when the binding of the naïve pool was examined, it also bound with a K_d of 1.3 μM; the two binding curves were superimposable. Thus, unlike standard aptamer selection in which binding function is necessary for selection, the regulatable, catalytically active nucleic acids of the present invention can be optimized for activation without affecting nascent binding. Given that lysozyme does not in general activate the random pool to any great degree this further emphasizes the specificity of the selected interface.

[00163] In natural ribonucleoproteins, protein components activate their nucleic acid counterparts by stabilizing active RNA conformers. The yeast mitochondrial protein CBP2 preferentially stabilizes the active tertiary structure of the intron bI5, while Cyt18 assists in folding and stabilization of the ND1 intron. The P-protein of RNase P has been shown to bind near the active site of the ribozyme and to influence substrate specificity. However, unlike ribonuclease P, the function of the protein cofactors of the present invention, nucleprotein enzymes cannot be replicated by simply increasing monovalent salt concentrations. Therefore the method of the present invention may be used to select regulatable, catalytically active nucleic acids in which activated catalysis is a synergistic property of the modified catalytic domain and its protein 'cofactor.' From this vantage, the role of the ribozyme would be to provide an adaptive platform for protein binding.

[00164] The ability to select ribozymes that are responsive to protein effectors has important implications for the development of biosensor arrays. The present invention may be used in conjunction with, or as a substitute for identifying antibodies to proteome targets, and are developing antibody-based chips for proteome analysis. However, the performance of such chips is inherently tied to the performance of antibodies. In order to develop sandwich-style assays, at least two different antibodies that recognize non-overlapping epitopes will need to be identified for each protein target, and the background binding of antibody:reporter conjugates will of necessity limit the sensitivity of ELISA-style assays. In contrast, protein-dependent regulatable, catalytically active nucleic acids could be immobilized on chips, transiently but specifically recognize their protein targets, covalently co-immobilize a reporter conjugated to an oligonucleotide substrate, and then be stringently washed to reduce background. The automation of *in vitro* selection procedures, as disclosed herein, demonstrate that it is possible to develop high-throughput regulatable, catalytically active nucleic acids selections, which could allow proteome and metabolome targets to be detected and quantitated.

[00165] Synthesis of L1-N50 pool and primers. The L1-N50 pool and primers were synthesized using standard phosphoramidite methodologies. Some 424 μg (ca. 10^{15} molecules) of the single stranded pool (5' TTCTAATACGACTCACTATAGGACCTCGGCGAAAGC-(N_{50})-GAGGTTAGGTGCCTCGTGATGTCCAGTCGC (SEQ ID NO:7) T7 promoter underlined, N = A, G, C, or T) was amplified in a 100 mL PCR reaction using the primers 20.T7 (5' TTCTAATACGACTCACTATA) (SEQ ID NO:8) and 18.90a (5' GCGACTGGACATCACGAG) (SEQ ID NO:9). The substrate used in the selection was S28A-biotin (biotin-$(dA)_{22}$-ugcacu; RNA in lowercase). A non-biotinylated version of this substrate (S28A) was used in most ligation assays. During selection, a selective PCR primer set, 28A.180 (5' $(dA)_{22}$-TGCACT)/18.90a, was used to amplify ligated ribozymes. A regenerative PCR primer set, 36.dB.2 (5' TTCTAATACGACTCACTATAGGACCTCGGCGAAAGC)(SEQ ID NO:10)/18.90a, restored the T7 promoter to the selected pool in preparation for further rounds of transcription and selection.

[00166] *In vitro* selection of protein dependent ribozymes. Briefly, pool RNA (5 μM) and 18.90a (10 μM) were first denatured in water. Ligation buffer (50 mM Tris, pH 7.5, 100 mM KCl, 10 mM $MgCl_2$) and substrate oligonucleotide (S28A-biotin, 10 μM) were then added in the absence of the target protein (except round 1). After this negative (-) incubation at 25°C, the selection mixture was segregated using a streptavidin-agarose resin (Fluka, Switzerland) to capture biotinylated substrate, free or ligated to the ribozyme. The eluant containing unligated ribozymes was collected and a second, positive (+) incubation was initiated by the addition of target protein (10 μM) and additional substrate (S28A-biotin, 10 μM). Following incubation at 25°C the mixture was again segregated using streptavidin-agarose. The resin containing ligated ribozymes was washed thoroughly and then suspended in RT buffer (50 mM Tris, pH 8.3, 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT, 400 μM dNTPs, 5 μM 18.90a) and reverse transcribed using SuperScript II reverse transcriptase (Gibco BRL, Gaithersburg, MD). The cDNA molecules in the resin slurry were then PCR amplified using first the selective primer set and then the regenerative primer set. The final PCR product was transcribed using T7 RNA polymerase (Epicentre, Madison, WI). Stringency was steadily increased over the course of the selection by decreasing the (positive selection) ligand incubation times and increasing the (negative selection) ligand incubation times (see Figs 18A and 18C).

[00167] Ligation assays. In one example, 10 pmol of [^{32}P]-body-labeled ribozyme and 20 pmol effector oligonucleotide were denatured for 3 minutes at 70°C in 5 μL water. The RNA mixture was cooled to room temperature followed by addition of ligation buffer and target protein (20 pmol unless otherwise stated, or water in place of ligand, in the case of (-) ligand samples). After a 5 minute equilibration at room temperature, reactions were initiated by the addition of 20 pmol substrate oligonucleotide (S28A) in a final volume of 15 μL. Reactions were incubated at 25°C, and 4 μL aliquots were removed at three appropriate time points and terminated by the addition of 18 μL of SDS stop mix (100 mM EDTA, 80% formamide, 0.8% SDS, 0.05% bromophenol blue, 0.05% xylene cyanol). Samples were denatured for 3 minutes at 70°C, ligated and unligated species were separated from one another on 8% polyacrylamide gels containing 0.1% SDS, and the amounts of products formed were determined using a Phosphorimager (Molecular Dynamics, Sunnyvale, CA).

Assays performed over a broad range of protein concentrations (e.g. Figure 21) differed from typical reaction conditions in that only 1 pmol ribozyme was present in a 10 μL final volume.

[00168] Protein inactivation. Standard ligation assays were performed as described above, but in the presence of protein samples that had been pre-treated as follows. Cyt18 protein was denatured by heating for 10 minutes at 70°C or by the addition of 6% SDS (0.7% SDS in ligation reaction). Lysozyme was heated 10 minutes at 100°C or incubated 10 minutes at room temperature in the presence of 2 mM DTT (0.3 mM DTT final reaction) without inactivating the protein. The protein was successfully inactivated by heating for 10 minutes at 70°C in the presence of 2 mM DTT. Ligation reactions were performed with 1.3 μM protein in 15 μL reactions incubated 5 minutes at 25°C.

[00169] Competition assays. Ligation assays were performed as described above, using 10 pmol of [^{32}P]-body-labeled ribozyme (cyt7-2 or lys11-2; 1 &M) and 20 pmol effector oligonucleotide (2 μM). The denatured and annealed RNA mixture was combined with ligation buffer, 20 pmol protein (Cyt18, lysozyme, or water in the case of (-) protein samples; 2 μM), and 30 pmol of denatured and annealed competitor RNA (3 μM). Competitor RNAs are as follows:

M12

[00170] GGGAA UGGAU CCACA UCUAC GAAUU CGAGU CGAGA ACUGG UGCGA AUGCG AGUAA GUUCA CUCCA GACUU GACGA AGCUU) (SEQ ID NO:11),

B17

[00171] GGGAA UGGAU CCACA UCUAC GAAUU CGUAG CGUAG AGUAU GAGAG AGCCA AGGUC AGGUU CACUC CAGAC UUGAC GAAGC UU) (SEQ ID NO:12)

c1

[00172] GGGAA UGGAU CCACA UCUAC GAAUU CAUCA GGGCU AAAGA GUGCA GAGUU ACUUA GUUCA CUCCA GACUU GACGA AGCUU (SEQ ID NO:13)

ND1

[00173] GACUA AUAUG AUUUG GUCUC AUUAA AGAUC ACAAA UUGCU GGAAA CUCCU UUGAG GCUAG GACAA UCAGC AAGGA AGUUA ACAUA UAAUG UUAAA ACCUU CAGAG ACUAG ACGUG AUCAU UUAAU AGACG CCUUG CGGCU CUUAU UAGAU AAGGU AUAGU CCAAA UUUGU AUGUA AAUAC AAAAU GAUAA AAAAA AAUGA AAUCA UAUGG G (SEQ ID NO:14)

N30

[00174] GGGAA UGGAU CCACA UCUAC GAAUU C-N30-U UCACU CCAGA CUUGA CGAAG CUU (SEQ ID NO:15)

[00175] Where N=(A, G, C, U), and tRNA (from Yeast; Gibco BRL, Gaithersburg, MD). Reactions were incubated 5 minutes at 25°C and initiated by the addition of 20 pmol substrate oligonucleotide (S28A; 2 μM) in a final volume of 10 μL. Cyt18 reactions were incubated 5 min at 25°C and lysozyme reactions were incubated 10 min. Reactions were terminated by the addition of 45 μL of SDS/urea stop mix (75 mM EDTA, 80% formamide, saturated urea, saturated SDS, 0.05% bromophenol blue, 0.05% xylene cyanol) and analyzed on 8% polyacrylamide gels containing 0.1% SDS as above.

[00176] Binding assays. Binding assays were performed in triplicate by combining 1 pmol of [^{32}P]-body-labeled RNA, 20 pmol 18.90a, and varying amounts of target protein (1 pmol to 5 nmol) in 50 μL of ligation buffer. After incubation at room temperature for 30 minutes, the mixture was drawn under vacuum through a series of nitrocellulose and nylon filters and washed with 150 μL of ligation buffer. The ratio of protein-bound RNA versus free RNA was determined by analyzing the counts retained on the nitrocellulose filter versus the counts on the nylon filter.

[00177] In Figure 17, L1 ligase, L1-N50 pool, and selection scheme. Figure 17 (a) shows the L1 ligase was the starting point for pool design. Stems A, B, and C are indicated. The shaded region indicates the catalytic core and ligation junction. Primer binding sites are shown in lower case, an oligonucleotide effector required for activity is shown in italics, and the ligation substrate is bolded. The 'tag' on the ligation substrate can be varied, but throughout this selection was biotin-$(dA)_{22}$. Figure 17 (b) shows the L1-N50 pool contains 50 random sequence positions and overlaps with a portion of the ribozyme core. Stem B was reduced in size and terminated with a stable GNRA tetraloop, and position U5 of stem A was mutated to a C (in bold) to form a base pair with G69 to increase the stability of the stem. Figure 17 (c) shows one selection scheme of the present invention. The RNA pool was incubated with a biotinylated substrate and reactive variants were removed from the population. The remaining species were again incubated with a biotinylated substrate in the presence of the target protein (Cyt18 or lysozyme). Reactive variants were removed from the population and preferentially amplified by reverse transcription, PCR, and *in vitro* transcription.

[00178] Figure 18 shows the progression of the L1-N50 selections. Figure 18(a) shows the conditions for the selection of Cyt18-dependent ribozymes. The 'substrate' column charts the molar excess of substrate to ribozyme. Figure 18(b) shows the progress of the L1-N50 Cyt18 selection. Ligation rates for each round of selection are plotted as black bars for assays performed in the presence of Cyt18 and gray bars for assays in the absence of Cyt18. The gray line the level of activation by Cyt18 and is measured against the right-hand axis. Figure 18(c) and 18(d) show the conditions for the selection of lysozyme-dependent ribozymes and the L1-N50 lysozyme selection. Graphing conventions are as in Figure 18b.

[00179] Figure 19 shows protein-dependent regulatable, catalytically active nucleic acid sequences and structures. Figure 19 (a) shows the sequences of the ribozyme N50 regions. Cyt18-dependent clones are indicated by the prefix 'cyt' and lysozyme-dependent clones are indicated by the prefix 'lys'. The number following these prefixes indicates the round from which the ribozyme was cloned (e.g., cyt7-2 was from the7th round of selection). The frequency that a given motif appears (out of 36 'cyt' clones and 24 'lys' clones) in the sequenced population is indicated in parentheses. Regions of sequence similarity between

individual clones are boxed. Figure 19(b) is a hypothetical secondary structure of the dominant Cyt18-dependent clone cyt7-2.

[00180] Figure 20 demonstrates the ribozyme activity with inactivated protein samples. Ligation assays for the Cyt18-dependent clone cyt9-18 and the lysozyme-dependent clone lys11-2 were performed in the presence of treated Cyt18 and lysozyme, respectively.

[00181] Figure 21 demonstrates an aptamer competition assays. Relative ligation activity of cyt7-2 and lys11-2 assayed in the presence of various specific and non-specific aptamer and RNA constructs. Samples labeled (+) contain activating protein with no competitor, while samples labeled (-) do not contain protein. The other samples contain either aptamers for Cyt18 (M12, B17) or lysozyme (c1), a group I intron that binds Cyt18 (ND1), or other non-specific RNAs as described in the text. Figure 21 shows the binding and ligation activity as a function of protein concentration. Fraction of lys11-2 RNA bound to lysozyme (open squares (G), left-hand axis) superimposed onto the reaction rate of lys11-2 RNA (closed circles (J), right-hand axis) over a range of lysozyme concentrations.

EXAMPLE 4

PEPTIDE SPECIFIC REGULATABLE, CATALYTICALLY ACTIVE NUCLEIC ACIDS

[00182] Rev-dependent RNA ligase ribozymes. An L1-N50 pool (10^{15} starting species) was subjected to an iterative regime of negative and positive selections for ligation activity. The pool was initially incubated with a biotinylated substrate and reactive species were removed; the pool was then mixed with the effector molecule, a 17 amino acid fragment of the HIV Rev protein, and reactive species were removed and amplified. The Rev peptide is a highly basic arginine rich motif whose natural function is as an RNA binding domain. In addition, RNA aptamers to the full Rev protein and the 17mer Rev peptide have been isolated using *in vitro* selection. During the course of the study the stringency of the negative selections was increased by increasing the time allowed for ligation and substrate

concentration in the absence of Rev peptide. The stringency of the positive selection step was increased by decreasing the time allowed for ligation and the substrate concentration.

[00183] Figure 22 is a flow chart of a method for negative and positive selection of RCANA according to the present invention. In step 10, the catalytic residues of a catalytic nucleic acid are identified. Next, a pool of oligonucleotides is generated in which at least one residue in the catalytic domain is mutated (step 12). In step 14, the pool of oligonucleotides is immobilized via, e.g., 3' hybridization to an affinity column followed by incubation of the immobilized oligonucleotide pool (step 16) with the cognate substrate of the catalytic residues. In the case of ligases, for example, those mutated pool members that maintain activity without the presence of an effector are removed from the pool (step 18). Step 18 is the negative selection step and the stringency may be increased or decreased by changing, e.g., the length of time of exposure between the enzyme and the ligand, salt and temperature conditions, buffers and the like. The remaining mutated members of the pool are incubated with an effector in step 20, which is the positive selection step for RCANA. The stringency of positive selection may also be affected by changing, e.g., the length of time of exposure between the enzyme and the ligand, salt and temperature conditions, buffers and the like. The pool members that become active, or more active, upon exposure to the effector in step 22 are removed, e.g., using capture ligases, the sequences are reverse transcribed in step 24 and isolated using, e.g., PCR using selective oligonucleotides for ligated species. These RCANA may be further selected and improved through subsequent rounds of selection, which may include the use of regenerative oligonucleotides that do not overlap the substrate binding portion of the RCANA followed by in vitro transcription and reintroduction into the system at, e.g., step 14.

round	(-) incubation substrate	(-) Cyt18	(+) incubation substrate	(+) Cyt18
1			2X	6 hr
2	2X	24 hr	2X	16 hr
3	2X	24 hr	2X	5 hr
4	2X	24 hr	2X	30 min
5	2X	48 hr	2X	5 min
6	2X	95 hr	2X	5 min
7	2X	95 hr	2X	1 min
8	2X	95 hr	2X	30 sec
9	5X	94 hr	2X	30 sec

[00184] The degree of peptide-dependent activation was assessed in a standard ligation assay. Ligation activity independent of the presence of Rev peptide progressively increased through Round 6 (Figure 24). By Round 7, the standard kinetic analysis of the population began to display two distinct phases indicating potentially that at least two different species of catalyst with different characteristics were becoming predominant in the population. The first phase indicated a population with fast ligation rate but which was not affected by the presence of peptide. The second phase indicated a population that was about 60-fold slower than the first phase population but which did show a small degree of peptide activation.

[00185] Two additional rounds of selection were performed with increased stringency in the negative selection and the final two rounds of the selection were cloned and sequenced. Kinetic analysis of the individual isolates revealed that the initial peptide-insensitive phase of the kinetic analysis could be contributed to a single clone (R8-1), which ligates with a fast rate (52 hr^{-1}) independent of the presence of peptide. Clone R8-1 is nearly identical to a ribozyme (JH1). A second clone (R8-4) showed Rev peptide induced activation. Clone R8-4 performed the ligation reaction with an observed rate of 0.86 h^{-1} in the presence of Rev peptide, but this rate dropped to 0.000046 h^{-1} when the peptide was left out of the reaction, a difference of 18,600-fold. Interestingly, the remaining four clones that were sequenced (including clone R8-2), which accounted for 65% of the final population, were completely inactive in the standard ligation assay. Additionally, when these clones were assayed in the presence of the round 9 pool RNA, ligation activity remained undetectable, eliminating the possibility that these clones are persisting in the population by using a parasitic trans-ligation

mechanism in which substrate is ligated onto these RNAs by some other ligase in the mixture in a trans-ligation reaction.

[00186] Specificity of activation. In order to assess the specificity of activation of selected ribozymes by peptide effectors, the Rev-dependent ligase was incubated with a variety of peptides, including HIV Tat, BIV Tat, bREX, bradykinin, as well as arginine. Activation was observed only with HIV Tat peptide at about 30%. In addition, the complete Rev protein was able to activate the ligase about 10% as well as the peptide. The ligase was assayed in the presence of different preparations of Rev peptide with different capping structures. All preparations of the Rev peptide activate the ligase but to slightly different extents. The selection was performed with a capped peptide (sREVn) that increases the degree of α-helicity of the peptide to mimic its conformation in the full Rev protein. A less capped peptide (aREV) with less a-helical character than sREVn was the best activator by about a factor of 2. These results suggest that activation is highly specific and not due to some contaminating factor (salt, magnesium) that might have been introduced during a particular peptide preparation. In addition, as several of the non-cognate peptides were known to bind RNA, both specifically and non-specifically, general stabilization of ribozyme structure by protein 'salts' was an unlikely mechanism for activation.

[00187] To further eliminate the possibility that some non-peptide contaminant of the peptide preparations was the actual activator of the ligase, the peptide was treated to destroy the peptide and then assayed to see if the sample could still activate the ligase. Peptide was treated with either a standard acid hydrolysis or a trypsin digestion. Neither treated peptide sample was able to activate the ribozyme.

[00188] Synthesis of L1-N50 pool and primers. The L1-N50 pool and primers were synthesized using standard phosphoramidite methodologies. Some 424 μg (ca. 10^{15} molecules) of the single stranded pool (5' TTCTAATACGACTCACTATAGGACCTCGGCGAAAGC-(N_{50})-GAGGTTAGGTGCCTCGTGATGTCCAGTCGC (SEQ ID NO:7) T7 promoter underlined, N = A, G, C, or T) was amplified in a 100 mL PCR reaction using the primers 20.T7 (5' TTCTAATACGACTCACTATA)(SEQ ID NO:8) and 18.90a (5'

GCGACTGGACATCACGAG)(SEQ ID NO:9). The substrate used in the selection was S28A-biotin (biotin-$(dA)_{22}$-ugcacu; RNA in lowercase). A non-biotinylated version of this substrate (S28A) was used in most ligation assays. During selection, a selective PCR primer set, 28A.180 (5' $(dA)_{22}$-TGCACT)/18.90a, was used to amplify ligated ribozymes. A regenerative PCR primer set, 36.dB.2 (5' TTCTAATACGACTCACTATAGGACCTCGGCGAAAGC)(SEQ ID NO:10)/18.90a, restored the T7 promoter to the selected pool in preparation for further rounds of transcription and selection.

[00189] *In vitro* selection of peptide dependent ribozymes. The selection procedure for protein dependent ligase ribozymes has been described herein above. Briefly, pool RNA (5 μM) and 18.90a (10 μM) were first denatured in water. Ligation buffer (50 mM Tris, pH 7.5, 100 mM KCl, 10 mM $MgCl_2$) and substrate oligonucleotide (S28A-biotin, 10 μM) were then added in the absence of the target protein (except round 1). After this negative (-) incubation at 25°C, the selection mixture was segregated using a streptavidin-agarose resin (Fluka, Switzerland) to capture biotinylated substrate, free or ligated to the ribozyme. The eluant containing unligated ribozymes was collected and a second, positive (+) incubation was initiated by the addition of target protein (10 μM) and additional substrate (S28A-biotin, 10 μM). Following incubation at 25°C the mixture was again segregated using streptavidin-agarose. The resin containing ligated ribozymes was washed thoroughly and then suspended in RT buffer (50 mM Tris, pH 8.3, 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT, 400 μM dNTPs, 5 μM 18.90a) and reverse transcribed using SuperScript II reverse transcriptase (Gibco BRL, Gaithersburg, MD). The cDNA molecules in the resin slurry were then PCR amplified using first the selective primer set and then the regenerative primer set. The final PCR product was transcribed using T7 RNA polymerase (Epicentre, Madison, WI). Stringency was steadily increased over the course of the selection by decreasing the ligand incubation times (positive selection) and increasing the ligand incubation times (negative selection) (see Table 1).

[00190] Ligation assays. Ligation assays were performed as described hereinabove. Typically, 10 pmol of [^{32}P]-body-labeled ribozyme and 20 pmol effector oligonucleotide were denatured for 3 minutes at 70°C in 5 μL water. The RNA mixture was cooled to room temperature followed by addition of ligation buffer and target peptide (20 pmol unless

otherwise stated, or water in place of ligand, in the case of (-) ligand samples). After a 5 minute equilibration at room temperature, reactions were initiated by the addition of 20 pmol substrate oligonucleotide (S28A) in a final volume of 15 μL. Reactions were incubated at 25°C, and 4 μL aliquots were removed at three appropriate time points and terminated by the addition of 18 μL of SDS stop mix (100 mM EDTA, 80% formamide, 0.8% SDS, 0.05% bromophenol blue, 0.05% xylene cyanol). Samples were denatured for 3 minutes at 70°C, ligated and unligated species were separated from one another on 8% polyacrylamide gels containing 0.1% SDS, and the amounts of products formed were determined using a Phosphorimager (Molecular Dynamics, Sunnyvale, CA). Assays performed over a broad range of peptide concentrations differed from typical reaction conditions in that only 1 pmol ribozyme was present in a 10 μL final volume.

[00191] Peptide inactivation. Standard ligation assays were performed as described above, but in the presence of peptide samples that had been pre-treated as follows. Peptide (15 nmol) was either hydrolyzed for 24 hours in 6 M HCl at 100°C or digested with trypsin-immobilized agarose resin 14 hours at 37°C. Both samples were evaporated to dryness and resuspended in water to a final concentration of 150 μM and used in place of peptide in standard ligation assays. In addition, control samples for hydrolysis and trypsin digestion containing no peptide were treated as described for peptide samples and tested to insure that they did not inhibit ligation in the presence of intact peptide.

[00192] Figure 23 shows the selection scheme for peptide binding. The RNA pool was incubated with a biotinylated substrate and reactive variants were removed from the population. The remaining species were again incubated with a biotinylated substrate in the presence of the target peptide. Reactive variants were removed from the population and preferentially amplified by reverse transcription, PCR, and *in vitro* transcription.

[00193] Figure 24 shows the progress of the L1-N50 Rev selection. Ligation rates for each round of selection are plotted as black bars for assays performed in the presence of Rev peptide and gray bars for assays in the absence of Rev peptide. The gray line indicates the level of activation by Rev peptide and is measured against the right-hand axis. The 'substrate' column charts the molar excess of substrate to ribozyme.

EXAMPLE 5

IN VIVO GENE REGULATION USING REGULATABLE, CATALYTICALLY ACTIVE NUCLEIC ACIDS

[00194] The present invention also includes the design and isolation of regulatable, catalytically active nucleic acids generated *in vitro* by design and selection for use *in vivo*. The regulatable, catalytically active nucleic acids disclosed herein permit the control of gene regulation or viral replication *in vivo*. The present inventors have generated regulatable, catalytically active nucleic acids that allows directed, *in vivo* splicing controlled by exogenously added small molecules. Substantial differences in gene regulation were observed with compounds that differed by as little as a single methyl group. Regulatable, catalytically active nucleic acids may find applications as genetic regulatory switches for generating conditional knockouts at the level of mRNA or for developing economically viable gene therapies.

[00195] In order to convert the Group I self-splicing intron into a regulatable, catalytically active nucleic acid, it was necessary to first identify sequences or structures in the catalytic domain of a ribozyme whose conformation might modulate splicing activity. One example of a ribozyme catalytic domain that may be used with the present invention is the Group I self-splicing intron because its structural and kinetic properties and interaction with the thymidylate synthase (td) gene in bacteriophage T4 have been extensively studied. A series of nested deletions of the P6 stem-loop partially or completely compromise ribozyme activity. More importantly, either magnesium or the tyrosyl tRNA synthetase from Neurospora mitochondria (CYT-18) can suppress many of these defects. Other introns have also revealed that deletion of the P5 stem-loop can modulate ribozyme activity. The present inventors recognized that sites where deletions modulated ribozyme activity might also prove to be sites where conformational changes to a nucleic acid may modulate catalytic activity. A series of Group I aptazymes were designed in which the anti-theophylline aptamer was substituted for either a portion of P6 or P5 (Figure 25). The point of attachment of the anti-theophylline sequence was selected for the design of theophylline-dependent cleavases and ligases.

[00196] The self-splicing activities of the Group I, regulatable, catalytically active nucleic acids were examined *in vitro* using a standard splicing assay. The stringency of ligand-induced suppressions of splicing defects was examined by carrying out the reactions at either low (3 mM, stringent) or high (8 mM, permissive) magnesium concentrations. Several of the constructs were inactive (e.g., Th3P6, Th5P6, and Th6P6) or showed no differential splicing activity (e.g., Th4P6 and Th2P5), but four constructs, Th1P6, Th2P6, Th3P6, and Th1P5, showed increased self-splicing in the presence of theophylline. For all of the nucleic acids except Th3P6, the ligand-induced splicing activity was greater in a standard assay at the more stringent magnesium concentration (see Table below).

[00197] The table below shows the relative *in vitro* splicing activity of constructs containing anti-theophylline aptamers. Extent of reaction is relative to the parental construct in 3 mM $MgCl_2$ with no theophylline at 2 hrs.

[$MgCl_2$]	3 mM		8 mM	
[Theo]	1.5 mM	0 mM	1.5 mM	0 mM
Parental	0.85	1.00	0.61	0.68
B11	0.03	0.02	0.31	0.34
Th1P6	0.05	0.20	0.31	0.16
Th2P6	0.04	0.15	0.31	0.04
Th3P6	0.03	0.04	0.20	0.04
Th4P6	0.05	0.06	0.38	0.37
Th5P6	0.04	0.00	0.05	0.03
Th6P6	0.03	0.01	0.00	0.03
Th1P5	1.08	0.91	0.85	0.74
Th2P5	0.70	0.57	0.03	0.03

[00198] The construct Th3P6 was inactive at lower magnesium concentrations, and the more permissive concentration was required to observe ligand-induced splicing activity. Interestingly, those constructs that showed ligand-dependent activity closely resembled the original deletion variants that showed magnesium-dependent recovery of splicing activity. For example, the junction between the binding and the Group I catalytic domain in the activatable regulatable, catalytically active nucleic acids Th2P6 resembled the construct td □P6-6 whose splicing defect at 3 mM magnesium was suppressed by 8 mM magnesium or by stabilization of the capping tetraloop sequence. Defects that poise a ribozyme between active and inactive conformers have previously been used to engineer effector-dependence.

[00199] Next, the extent of ligand-dependent activation was determined by examining the kinetics of splicing in the presence and absence of theophylline. The nucleic acid modified at P5 (Th1P5) showed very little (1.6-fold) activation. Nucleic acids modified at P6 showed somewhat greater activation, with Th2P6 yielding 9-fold activation and Th1P6 18-fold initial rate enhancement in the presence of theophylline (data not shown). These levels of ligand-dependent activation were similar to those observed with the hammerhead ribozyme constructs, and it may prove possible to use *in vitro* selection to further optimize activation using the materials and methods of the present invention.

[00200] The mechanism of activation on the nucleic acids disclosed herein is likely the same as has been observed for other nucleic acids: ligand-induced conformational changes that stabilize functional nucleic acid sequences and structures. However, the Group I self-splicing intron is a much more complicated ribozyme than either the hammerhead or the L1 ligase; for example, the tertiary structure of the Group I intron is established by a complicated folding pathway. Therefore, it was possible that theophylline-binding influenced the overall folding or stability of the engineered Group I aptazyme, rather than merely altering the local conformation of a functional structure. In order to assess this possibility the theophylline-dependence of splicing reactions *in vitro* was examined following prolonged incubation to allow re-folding and initiation of catalysis with exogenous GTP. No change in the degree or rate of ligand-dependent activation was observed following pre-incubation (data not shown). Similarly, when theophylline was added to an *in vitro* splicing reaction that had previously been initiated with GTP, an increase in the rate of splicing to levels previously observed in the presence of theophylline was observed (data not shown). Taken together, these results militate against the assumption that theophylline influences the folding pathway of the engineered Group I aptazymes.

[00201] An attempt was made to change the effector specificity of the Group I aptazyme by changing which aptamer sequence was conjoined to the catalytic core. Previous studies with both the native hammerhead ribozyme and the L1 ligase showed that such swaps of allosteric binding sites and effector specificities were frequently possible. Soukup, G. A. & Breaker, R. R. Engineering precision RNA molecular switches. *Proc. Natl. Acad. Sci. U. S. A.* 96, 3584-3589 (1999), and Robertson, M. P. & Ellington, A. D. Design and

optimization of effector-activated ribozyme ligases. *Nucleic Acids Res.* 28, 1751-1759 (2000). To this end, the two most successful P6 constructs, Th1P6 and Th2P6, were re-engineered so that the anti-FMN aptamer was inserted in place of the anti-theophylline aptamer. The point of attachment of the anti-FMN aptamer was the same as had previously proven successful in the design of other FMN-dependent ribozymes (Figure 26). Both flavin-sensing Group I aptazymes were activated by FMN in a standard assay as well as or better than the theophylline-sensing Group I aptazymes. This result is especially significant given that FMN inhibits Group I splicing activity (albeit at concentrations higher than disclosed herein). Similar specificity swaps were attempted with anti-ATP and anti-HIV-1 Rev binding sequences, but neither of these potential allosteric binding sites appeared to communicate with the catalytic core of the intron. The anti-FMN aptamer may have been more readily substituted for the anti-theophylline aptamer because both terminate in an A:G base-pair. It may be that a different connecting stem or 'communication module' would allow the melding of other allosteric domains with the Group I ribozyme.

[00202] In the table below, the relative *in vitro* splicing activity of constructs containing anti-FMN aptamers is shown. The extent of reaction is relative to the parental construct in 3 mM $MgCl_2$ with no FMN at 2 hrs.

[$MgCl_2$]	3 mM		8 mM	
[FMN]	1 mM	0 mM	1 mM	0 mM
Parental	0.84	1.00	0.89	0.79
B11	0.14	0.05	0.08	0.50
FMN1P6	0.08	0.61	0.56	0.65
FMN2P6	0.06	0.41	0.44	0.19

[00203] Each of the successful nucleic acid constructs disclosed herein was subsequently cloned into an interrupted thymidylate synthetase gene in place of the parental td self-splicing intron. The vectors were introduced into an *E. coli* strain (C600ThyA::Kan^R) that lacked a functional thymidylate synthetase gene and that were thymidine auxotrophs. When bacteria grown in rich media were subsequently plated on minimal media lacking thymidine, no colony growth was observed with the exception of Th1P5. However, when theophylline (7.5 mM) was included in the minimal media, colony growth was observed for the intron Th2P6 and increased growth for Th1P5 (data not shown). Interestingly, no growth

was observed for constructs harboring the intron Th1P6, despite the fact that this nucleic acid showed a much greater level of theophylline-enhanced splicing *in vitro*. All introns that originally showed no or low splicing *in vitro* (including Th3P6) could not rescue cells either in the presence or absence of theophylline. Finally, no growth was observed in a negative control that contained a non-functional Group I intron (B11) and no growth change in a negative control in which mutations were introduced to abolish theophylline binding (Th2P6.D) either in the presence or absence of theophylline.

[00204] To better quantitate the influence of the effector on intron-splicing, growth experiments in liquid culture were conducted (Figure 27(a)). An overnight culture that contained the td gene divided by the nucleic acid Th2P6 was inoculated into fresh, minimal media, effector was added, and the resultant growth curves were continuously monitored. As expected based on the results from growth assays on solid media, little growth is observed in the absence of theophylline. However, when theophylline (0.5 mM) is added to liquid medium, cells grow almost as well as a control in which the parental intron is inserted into the td gene.

[00205] Importantly, cell growth is not activated by the structurally-related effector caffeine (i.e., 7-methyltheophylline), and no effector-dependent growth differences are observed with cultures containing td genes divided by the non-functional Group I intron B11. The anti-theophylline aptamer is known to discriminate between caffeine and theophylline by a factor of 10,000-fold. Similar results were obtained with cultures that contained the td gene divided by the nucleic acid Th1P5 (Figure 27(b)). However, in this instance there was some background growth of uninduced cells, consistent with the higher level of background splicing activity *in vitro*. If theophylline is regulating intron splicing *in vivo*, then the extent of cell growth should be dependent upon the concentration of theophylline introduced into the media (Figure 27(c)). Theophylline was toxic to cells, and caused a decrease in the growth of cells containing the parental td intron at concentrations greater than 0.5 mM. Low concentrations of theophylline progressively increase cell growth (by activating the td intron) while concentrations of theophylline above 2 mM progressively decrease cell growth (although levels of growth are still well above background).

[00206] The presence of endogenous flavins made it difficult to examine effector-specificity *in vivo*, and a new series regulatable, catalytically active nucleic acids were constructed in which the anti-theophylline binding sequence was mutated to bind 3-methylxanthine (3MeX2P6). These variants proved to be responsive to 3-methylxanthine both *in vitro* and *in vivo* (Figure 28). However, the variants were no longer responsive to theophylline, nor were they responsive to a variety of other analogues, including caffeine, 1-methylxanthine, 7-methylxanthine, 1,3-dimethyl urilic acid, hypoxanthine, xanthine, and theobromine (data not shown).

[00207] These results indicate that theophylline regulates intron splicing *in vivo*. Next, mRNA was isolated from E. coli treated in the presence or absence of theophylline, and RT-PCR was used to confirm the presence of spliced introns (data not shown). For each of the introns known to be responsive to theophylline *in vivo* (Th2P6 and Th1P5) an increase in spliced mRNA is observed, while those introns not responsive to theophylline *in vivo* did not show an increase in the levels of spliced mRNA. An exception to this was Th1P6, which originally showed theophylline-dependent splicing *in vitro* and theophylline-dependent splicing *in vivo*. However, Th1P6 does not mediate theophylline-dependent growth. The cellular mRNAs were extracted, cloned, and sequenced, and half of them appeared to use a cryptic splice site.

[00208] The ability to engineer regulatable, catalytically active nucleic acids to be responsive to effector molecules has numerous potential applications. For example, it may be used in conjunction with new gene therapies in which patients rely upon drugs that differentially activate gene expression, rather than having to rely upon a set level of endogenous expression of an introduced gene. Similarly, it may be used with effector-dependent splicing to more finely monitor gene expression *in vivo*. A drug that localized to particular organs, cells, or organelles, and splicing of the nucleic acid could be monitored via a reporter gene such as, e.g., luciferase. Engineered introns introduced into reporter genes may be used in high-throughput, cell-based screening assays that monitor drug uptake or efficacy.

[00209] Materials and Methods. *E. coli* strains and growth media. *E. coli* strain C600ThyA::KanR was used for the plate assays and *in vivo* growth curves. INVaF' (Invitrogen, Carlsbad, CA) was used for cloning and plasmid amplification. Bacterial starter cultures were grown in LB supplemented with thymine (50 mg/l). Screening for the td phenotype was done in minimal media supplemented with 0.1% Norit A-treated casamino acids (MM) and MM supplemented with thymine (50 mg/l) (MMT). Plates contained Bacto agar (1.5%). Ampicillin (50 mg/l) and kanamycin (100 mg/l) were added to all growth media.

[00210] Plasmid. The wild type plasmid pTZtd1304 (Myers et al 1996) contains a 265 nucleotide derivative of the 1016 nucleotide wild type intron that maintains wild type activity (Galloway Salvo et al 1990) with additional mutations of U34A which introduces a SpeI site and U976G which introduces an EcoRI site.

[00211] Construction of the *td* intron regulatable, catalytically active nucleic acids. The constructs were made using standard solid phase DNA synthesis, then were PCR-amplified and cloned into pTZtd1304 that contained a 265 nucleotide derivative of the 1016 nucleotide wild-type intron. This derivative also contained the mutations U34A, which introduces a SpeI site, and U976G, which introduces an EcoRI site. The parental P6 nucleic acid construct was generated by two overlapping oligos, Gp1Wt2 Gp1Wt2.122 (GCC TGA GTA TAA GGT GAC TTA TAC TTG TAA TCT ATC TAA ACG GGG AAC CTC TCT AGT AGA CAA TCC CGT GCT AAA TTG TAG GAC TGC CCG GGT TCT ACA TAA ATG CCT AAC GAC TAT CCC TT) (SEQ ID NO:16); and

[00212] Gp1Wt3.129 (TAA TCT TAC CCC GGA ATT ATA TCC AGC TGC ATG TCA CCA TGC AGA GCA GAC TAT ATC TCC AAC TTG TTA AAG CAA GTT GTC TAT CGT TTC GAG TCA CTT GAC CCT ACT CCC CAA AGG GAT AGT CGT TAG) (SEQ ID NO:17). These oligonucleotides (100 pmol) were annealed and extended with AMV reverse transcriptase (Amersham Pharmacia Biotech, Piscataway, NJ; 45 units) in AMV RT buffer (50 mM Tris-HCl, pH 8.3, 8 mM $MgCl_2$, 50 mM NaCl, 1 mM DTT) and dNTPs (200 μM) for 30 minutes at 37° C. The resulting double-stranded DNA was diluted 1:50 and amplified using primers SpeI.24 (TTA TAC TAG TAA TCT ATC TAA ACG

(SEQ ID NO:18); 0.4 μM) and EcoRI.24 (CCC GGA ATT CTA TCC AGC TGC ATG (SEQ ID NO:19); 0.4 μM) in PCR buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.1% Triton X-100, 0.005% gelatin), dNTPs (200 μM) and Taq DNA polymerase (Promega, Madison, WI; 1.5 units). The reactions were thermocycled 15 times at 94° C for 30 seconds, 45° C for 30 seconds, 72° C for 1 minute and then purified with a QIAquick PCR purification kit (Qiagen, Valencia, CA).

[00213] The PCR product was digested with SpeI (New England Biolabs, Beverly, MA; 20 units) and EcoRI (50 units) in buffer (50 mM NaCl, 100 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 0.025% Triton X-100, 100 μg/ml BSA) at 37° C for 60 minutes, purified, and cloned into SpeI/EcoRI digested pTZtd1304. The negative control and nucleic acid constructs were made as described except that Gp1Wt3.129 was replaced with oligonucleotides of the appropriate sequence: B11 GCC TGA GTA TAA GGT GAC TTA TAC TTG TAA TCT ATC TAA ACG GGG AAC CTC TCT AGT AGA CAA TCC CGT GCT AAA TGC CTA ACG ACT ATC CCT T (SEQ ID NO:20),

[00214] Th1P6 GCC TGA GTA TAA GGT GAC TTA TAC TTG TAA TCT ATC TAA ACG GGG AAC CTC TCT AGT AGA CAA TCC CGT GCT AAA TTA TAC CAG CAT CGT CTT GAT GCC CTT GGC AGA TAA ATG CCT AAC GAC TAT CCC TT (SEQ ID NO:21),

[00215] Th2P6 GCC TGA GTA TAA GGT GAC TTA TAC TTG TAA TCT ATC TAA ACG GGG AAC CTC TCT AGT AGA CAA TCC CGT GCT AAA TTG ATA CCA GCA TCG TCT TGA TGC CCT TGG CAG CAT AAA TGC CTA ACG ACT ATC CCT T (SEQ ID NO:22),

[00216] Th3P6 GCC TGA GTA TAA GGT GAC TTA TAC TTG TAA TCT ATC TAA ACG GGG AAC CTC TCT AGT AGA CAA TCC CGT GCA TAC CAG CAT CGT CTT GAT GCC CTT GGC AGG CCT AAC GAC TAT CCC TT (SEQ ID NO:23),

[00217] Th4P6 GCC TGA GTA TAA GGT GAC TTA TAC TTG TAA TCT ATC TAA ACG GGG AAC CTC TCT AGT AGA CAA TCC CGT GCT AAA TAT ACC AGC

ATC GTC TTG ATG CCC TTG GCA GTA AAT GCC TAA CGA CTA TCC CTT (SEQ ID NO:24),

[00218] Th5P6 GCC TGA GTA TAA GGT GAC TTA TAC TTG TAA TCT ATC TAA ACG GGG AAC CTC TCT AGT AGA CAA TCC CGT ATA CCA GCA TCG TCT TGA TGC CCT TGG CAG CTA ACG ACT ATC CCT T (SEQ ID NO:26),

[00219] Th6P6 GCC TGA GTA TAA GGT GAC TTA TAC TTG TAA TCT ATC TAA ACG GGG AAC CTC TCT AGT AGA CAA TCC CGT GAT ACC AGC ATC GTC TTG ATG CCC TTG GCA GCC TAA CGA CTA TCC CTT (SEQ ID NO:27),

[00220] Th1P5 TGA GTA TAA GGT GAC TTA TAC TAG TAA TCT ATC TAA ACG GGG AAC CTC TAT ACC AGC ATC GTC TTG ATG CCC TTG GCA GAG ACA ATC CCG TGC TAA ATT GTA GGA CTG CCC GGG TTC TAC ATA AAT GCC TAA CGA CTA TCC CTT (SEQ ID NO:28),

[00221] Th2P5 TGA GTA TAA GGT GAC TTA TAC TAG TAA TCT ATC TAA ACG GGG AAC CTA TAC CAG CAT CGT CTT GAT GCC CTT GGC AGA CAA TCC CGT GCT AAA TTG TAG GAC TGC CCG GGT TCT ACA TAA ATG CCT AAC GAC TAT CCC TT (SEQ ID NO:29),

[00222] 3Mex2P6 GTA ATC TAT CTA AAC GGG GAA CCT CTC TAG TAG ACA ATC CCG TGC TAA ATT GAT ACC AGC ATC GTC TTG ATG CCA TTG GCA GCA TAA ATG CCT AAC GAC TAT CCC TT (SEQ ID NO:30),

[00223] Th2P6.D GTA ATC TAT CTA AAC GGG GAA CCT CTC TAG TAG ACA ATC CCG TGC TAA ATT GAT ACC AGC ATC GTC TTG ATG CCC TTG GTT GCA TAA ATG CCT AAC GAC TAT CCC TT (SEQ ID NO:31),

[00224] FMN1P6 GCC TGA GTA TAA GGT GAC TTA TAC TTG TAA TCT ATC TAA ACG GGG AAC CTC TCT AGT AGA CAA TCC CGT GCT AAA TTA GGA TAT GCT TCG GCA GAA GGA TAA ATG CCT AAC GAC TAT CCC TT (SEQ ID NO:32), and

[00225] FMN2P6 GCC TGA GTA TAA GGT GAC TTA TAC TTG TAA TCT ATC TAA ACG GGG AAC CTC TCT AGT AGA CAA TCC CGT GCT AAA TTG AGG ATA TGC TTC GGC AGA AGG CAT AAA TGC CTA ACG ACT ATC CCT T (SEQ ID NO:33).

[00226] *In vitro* transcription. The introns were PCR-amplified with 5' le (GAT AAT ACG ACT CAC TAT AAT GGC ATT ACC GCC TTG T) (SEQ ID NO:34) and GM24 (GCT CTA GAC TTA GCT ACA ATA TGA AC) (SEQ ID NO:35) in 25 µl reactions under the conditions stated above and cycled 20 times. A portion of the reaction (5 µl) was run on a 3% agarose gel and the PCR product band was stabbed with a pipette tip. The agarose plug was added to a fresh PCR reaction (100 µl) and cycled 15 times; DNA was purified using a QIAquick kit and quantitated. The PCR product (2 µg in 50 µl) was added to an *in vitro* transcription reaction containing Ampliscribe T7 RNA polymerase (Epicentre), RNase inhibitor (GIBCO BRL, Rockville, MD; 5 units), low Mg2+ buffer (30 mM Tris-HCl, pH 8, 7.5 mM DTT, 4.5 mM MgCl2, 1.5 mM spermidine), UTP (1.25 mM), ATP (2.5 mM), GTP (2.5 mM), CTP (7.5 mM) and aP32-labeled UTP (NEN, Boston, MA; 20 µCi; 3000 mCi/mmol), and incubated at 37° C for 2 hours. DNase (GIBCO BRL, 295 units) was added and the reaction was incubated at 37° C for an additional 30 minutes. The RNA was purified using Centri-Sep columns (Princeton Separations, Adelphia, NJ) and quantitated.

[00227] *In vitro* splicing assays. The assays were preformed by heating the RNA (500 nM) in H_20 to 70° C for 3 minutes then transferring to ice for 1 minute. Splicing buffer (20 mM Tris-HCl, pH 7.5, 100 mM KCl, 3 mM $MgCl_2$), effector (Theophylline (1.5 mM) or FMN (1 mM)) was added and the reactions were incubated on ice for an additional 15 minutes. At this time a 4.5 µl aliquot was removed for a zero time point and quenched with 5 µl stop dye (95% formamide, 20 mM EDTA, 0.5% xylene cyanol, and 0.5% bromophenol blue). GTP (50 µM) was added to the remaining reaction (5 µl total volume) to start the splicing reaction. The reaction was incubated at 37° C for 30 minutes and then terminated with stop dye (5 µl). The reactions were heated to 70° C for 3 minutes and 5 µl was analyzed on an 8% denaturing polyacrylamide gel. The gel was dried, exposed to a phosphor screen and analyzed using a Molecular Dynamics Phosphorimager (Sunnyvale, CA).

[00228] The reaction volumes were increased for the rate determination assay. Aliquots were taken at intervals between 0 minutes and 30 minutes and terminated with stop dye. The reactions were analyzed as described above.

[00229] *In vivo* plate assay. The plasmids containing the various group I constructs were transformed into chemically competent C600ThyA::Kan^R cells and grown in LB with kanamycin overnight. A small aliquot (3 µl) of overnight cell culture was mixed with effector (theophylline (7.5 mM) or FMN (10 mM)) or H_20, spotted on plates, and grown overnight at 37 °C. As a positive control, all constructs were also plated on minimal media plates with thymine (MMT) and assayed for viability.

[00230] *In vivo* growth curves. Cells grown overnight in LB were diluted 1:100 in MM containing either theophylline, caffeine, 3-methylxanthine or no effector, and analyzed on a Microbiology Workstation Bioscreen C (Labsystems, Inc., Franklin, MA).

[00231] RT-PCR analysis. RNA was isolated from an overnight culture using a MasterPure RNA purification kit (Epicentre, Madison, WI) and amplified by RT-PCR using primers 5' le and GM24 following the protocol provided for Tth polymerase. The products were separated and analyzed on a 3% agarose gel.

[00232] Figure 25 shows the theophylline-dependent *td* group I intron constructs of the present invention. The Figure 25(a) shows the predicted secondary structure and tertiary interactions of the 265 nucleotide deletion construct of the *td* intron. The intron is in uppercase and the exons are in lower case letters. The 5' and 3' splice sites are indicated by arrows. The P4-P6 domain is boxed. Figure 25(b) shows the B11 construct based on the Δ85-863 deletion mutant of the td intron, which shows no activity at low Mg^{2+} (3 mM) *in vitro* or *in vivo*. An anti-theophylline aptamer, highlighted in gray, was substituted for the P6a stem of the intron in constructs Th1P6, Th2P6, Th3P6, Th4P6, Th5P6 and Th6P6, and for the P5 stem in constructs Th1P5 and Th2P5. Mutations in the anti-theophylline aptamer are boxed in black for constructs MeX2P6 and Th2P6.D. The C-to-A mutation in MeX2P6 changes specificity from theophylline to 3-methylxanthine. The A-to-U and C-to-U mutations in Th2P6.D abolished theophylline-binding.

[00233] The *in vitro* activation of td group I nucleic acids by theophylline was also demonstrated (data not shown). The splicing activity of the parental, B11, Th1P6, Th2P6 and Th1P5 intron constructs in the presence and absence of 1.5 mM theophylline using autoradiography in which the following products were identified: LI, linear intron; CI, circular intron; E1–E2, exon 1–exon 2 ligation product; Crp, cryptic ligation product; pre-mRNA, unspliced mRNA (data not shown).

[00234] Figure 26 shows the design of an FMN-dependent *td* nucleic acid intron splicing construct. An anti-FMN aptamer, highlighted in gray, was substituted for the P6a stem in constructs FMN1P6 and FMN2P6. *In vivo* splicing activity was demonstrated on agar plates. The parental, B11 and theophylline constructs were spotted in the presence and absence of 7.5 mM theophylline on minimal media (MM), while the parental, B11 and FMN constructs were spotted in the presence and absence of 5 mM FMN (data not shown).

[00235] Theophylline-dependent *in vivo* growth was assayed and quantitated. Figures 27(a), 27(b) and 27(c) show the relative growth curves are shown for C600:ThyA cells containing either Th2P6 (a) and Th1P5 (b) in the presence (□) and absence (□) of 0.5 mM theophylline or 0.5 mM caffeine (□). Parental (□) and B11 (□) controls were grown in the 0.5 mM theophylline for comparison. Plots are standardized to the growth of cells containing the parental intron. Each point represents the average of three replicate growth curves. Figure 27(c) shows the extent of growth at 12 hours for parental, Th2P6 and Th1P5 introns over a range of theophylline concentrations. Background growth (B11) has been subtracted, and results are standardized to parental growth with no theophylline.

[00236] Figure 28 shows the 3-Methylxanthine dependent *in vivo* growth. Relative growth curves are shown for C600:ThyA cells containing 3MeX2P6 in the presence (□) and absence (□) of 1 mM 3-methlyxanthine or 1 mM theophylline (□). Parental (□) and B11 (□) controls were also grown in 1 mM 3-methylxanthine. Plots are standardized to parental growth. Each point represents the average of three replicate growth curves. To shows the splicing of introns *in vivo*, RT-PCR analysis of whole cell RNA was conducted. Bands corresponding to spliced and unspliced mRNAs were identified (data not shown). Samples

was seeded with RNA from cells grown in the absence of theophylline and compared with samples seeded with RNA from cells grown in the presence of 0.5 mM theophylline.

EXAMPLE 6

DETECTION OF A DIVERSE SET OF ANALYTES USING ARRAYED RIBOZYME LIGASES

[00237] Several catalytic RNAs have been shown to be amenable to engineering. In several cases, a particular ribozyme scaffold can be evolved and engineered to respond to a wide variety of effectors. These properties give regulatable, catalytically active nucleic acids, tremendous potential in the field of molecular diagnostics. The engineering of the hammerhead ribozyme can yield variants that are allosterically regulated by a variety of ligands (Koizumi, M.; Kerr, J. N.; Soukup, G. A.; Breaker, R. R. *Nucleic Acids Symp Ser.*, 1999, 42, 275-27). In addition, several of these allosteric hammerhead variants have in turn been used to assemble a ribozyme array able to detect a variety of small-molecules.

[00238] In order to demonstrate the utility of ribozyme ligases in multiplexed, multiple analyte assays, a series of ligases previously developed by the inventors (described hereinabove) were used in an array. Notably, the array can detect a diverse range of biologically relevant analytes: small-molecules, nucleic acid, a protein and a peptide may be assayed in solution.

[00239] Regulatable ligase variants were evolved starting with a small ribozyme ligase, L1, which was initially selected from a random sequence pool. The activity of this ribozyme was found to be dependent upon the 3' primer used in the selection, increasing the ribozyme's activity up to 10,000 fold in its presence. Additional L1 variants have been designed or selected to respond to small-molecules (ATP, FMN, theophylline), proteins (lysozyme), and peptides (Rev).

[00240] As an initial test of the ability of this ensemble of regulatable, catalytically active nucleic acids to function in a multiplexed assay, a simple scheme was developed for monitoring the self-attachment of the ligases to 96-well plates. By virtue of a biotinylated

substrate, ligation of radio-labeled ribozymes in response to a given analyte can be monitored by quantitating the fraction immobilized in streptavidin coated polystyrene plates (Figure 29).

[00241] A typical regulatable, catalytically active ligase array is depicted in Figure 30. All the regulatable, catalytically active nucleic acids used (rows) were tested against the corresponding set of ligands (columns). The diagonal represents a positive reaction between an regulatable, catalytically active nucleic acids and its cognate ligand. All regulatable, catalytically active nucleic acids were also tested for activity in complex mixtures ('+' column, mixture of all 6 ligands), as well as inactivity in the absence of effector ('-' column). For the most part, there is extremely high specificity between a particular regulatable, catalytically active nucleic acids and its cognate ligand. All of the regulatable, catalytically active nucleic acids retained activity in the context of a complex mixture. Note the cross-reactivity of L1-ATP with flavin mononucleotide (FMN), which may be due to chemical similarity between FMN and ATP. The array depicted in Figure 30 is the 'positive' image of a typical assay; the supernatant removed following an assay was transferred to a separate plate for the quantitation of background and unligated species.

[00242] In order to better characterize individual aptazymes' properties in the context of an array, their ability to carry out ligation to a plate-bound substrate was monitored in response to ligand concentration (Figure 31). Aptazymes (rows) were assayed in array format against the corresponding set of analytes (columns). Many of the aptazyme's activities are similar to values calculated previously. All of the ribozymes assayed displayed response characteristics with Kd's in the high nM to low μM range.

[00243] Figure 29 shows a schematic of ribozyme ligase array. In 29(a) the absence of analyte, the ribozyme is unable to catalyze the ligation of biotinylated substrate, and remains in the supernatant. In Figure 29(b) analyte concentrations high enough to cause ligation result in the self-attachment of a tagged substrate, which is then immobilized to streptavidin-coated 96-well plates.

[00244] Figure 30 shows the results of a regulatable, catalytically active ligase array. Regulatable, catalytically active nucleic acids and effector pairs are assayed in array format;

the 'positive' plate is pictured. The diagonal represents a positive reaction between a ribozyme and its cognate ligand.

[00245] Figure 31 shows the titrations of individual allosteric ribozyme ligases. Response curves for five individual aptazymes are calculated. Normalized counts are plotted against cognate effector concentration (e.g. L1-FMN activity vs. [FMN]). Kd's are calculated by fitting data to a simple saturation curve (y=(m1*m0)/(Kd+m0)). The maximum percentage bound to the 'positive' plate is reported to illustrate the extent of ligation over the time allotted.

[00246] Sequences. Sequences for L1, L1-ATP, L1-FMN, and L1-theophylline have been published previously, while L1-Rev was recently selected: (SEQ). The 5' primer used in PCR amplification incorporates a T7 promoter, while the 3' primer is universal for all templates.

[00247] RNA Preparation. Individual ribozymes were generated by standard *in vitro* transcription reactions containing 500ng of PCR product, Tris-HCl, DTT, each of the four ribonucleotides, and 50 U of T7 RNA Polymerase. Following gel purification, the RNAs were eluted in water, precipitated and resuspended in water.

[00248] Aptazyme Array and Titration of Individual Aptazymes. Arrayed aptazyme assay were carried out by first annealing 100 pmol of ribozyme with 120 pmol of 18.90A (5' GCGACTGGACATCACGAG 3')(SEQ ID NO:36). Following addition of buffer (30 mM Tris-HCl, pH 7.5, 50 mM NaCl, 60 mM $MgCl_2$), 120 pmol of substrate (S28A-biotin, 5' biotin-AAAAAAAAAAAAAAAAAAAAAAugcacu 3', (SEQ ID NO:37) ribonucleotides in lowercase) was added. The reaction mixture was scaled to accommodate multiple aliquots for each corresponding well of the array. After aliquotting 50 µl into each well of an 96-well PCR plate (MJ Research), 50 µl of ligand in buffer was added. Ligand concentrations for Fig. 29 were: 1 µM 18.90A, 0.5 mM flavin mononucleotide (FMN), 5 µM lysozyme, 1 µM Rev peptide, 1 mM ATP, and 1 mM theophylline.

[00249] Reactions were incubated at 25°C for 4 hours, followed by the addition of 20 µl of 0.5 M EDTA. Reactions were then transferred to Hi-Bind streptavidin coated

polystyrene plates (Pierce). Plates were again incubated at room temperature for one hour, followed by the transfer of supernatant to a plain polystyrene 96-well plate. Wells in the Hi-bind plates were washed three times with buffer (30 mM Tris-HCl pH 7.5, 100 mM NaCl, 0.1% SDS, 7 M urea), followed by a rinse in TE (10 mM Tris-HCl, pH 7.5, 1 mM EDTA). Assays were quantitated by exposure to PhophorImager plates followed by analysis with ImageQuant software (Molecular Dynamics). Titrations (Fig. 31) were carried out essentially as described previously, with ligand titrated in a range a concentration.

[00250] All publications mentioned in the above specification are hereby incorporated by reference. Modifications and variations of the described compositions and methods of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described compositions and modes of carrying out the invention that are obvious to those skilled in molecular biology or related arts are intended to be within the scope of the following claims.